



Universidad de Granada

**EXTRACCIÓN EFICIENTE DE LA
ESTRUCTURA DE ESCENAS NATURALES**

JOSÉ MANUEL PALOMARES MUÑOZ

Para optar al grado de
DOCTOR EN INFORMÁTICA

noviembre 2010

Editor: Editorial de la Universidad de Granada
Autor: José Manuel Palomares Muñoz
D.L.: GR 1862-2011
ISBN: 978-84-694-1304-3

VISTO BUENO

El **Prof. Dr. D. Jesús González Peñalver**, Profesor Titular de Universidad y el **Prof. Dr. D. Eduardo Ros Vidal**, Catedrático de Universidad, ambos del Departamento de Arquitectura y Tecnología de Computadores de la Universidad de Granada,

CERTIFICAN:

Que la memoria titulada:

“Extracción Eficiente de la Estructura de Escenas Naturales ”

ha sido realizada por **D. José Manuel Palomares Muñoz** bajo nuestra dirección en el Departamento de Arquitectura y Tecnología de Computadores de la Universidad de Granada para optar al grado de **Doctor en Informática** .

En Granada, a 14 de noviembre 2010.

Los Directores de la tesis doctoral:

Fdo. Jesús González Peñalver

Fdo. Eduardo Ros Vidal

A mi familia:
Cristina, Pablo y Pili.
Paco, Ana y Ana María.

RESUMEN

Esta Tesis Doctoral pretende proporcionar un mecanismo que permita la extracción eficiente de las características presentes en una escena natural para pacientes con BAJA VISIÓN. Con esta motivación, se ha desarrollado este trabajo, que presenta dos aportaciones científicas principales. La primera es la adaptación del operador *convolución* bajo el paradigma LIP, de tal forma que ésta permite la construcción de métodos basados en filtros separables en dicho paradigma. Debido al carácter logarítmico de LIP, gracias a este operador se pueden diseñar algoritmos de extracción de contornos que son invariantes ante cambios de iluminación. A lo largo de la Tesis Doctoral se adaptan diversos algoritmos de extracción de bordes tradicionales al paradigma LIP utilizando el operador propuesto.

La segunda aportación es la evaluación SUBJETIVA mediante cuestionarios de la calidad de los mapas de contornos obtenidos por cada método. En esta Tesis Doctoral, se proponen una serie de recomendaciones para la construcción de buenas encuestas para la evaluación de imágenes de contornos, que han sido utilizadas para evaluar la calidad de los métodos adaptados al paradigma LIP utilizando el nuevo operador propuesto. La muestra poblacional que responde a dicha encuesta incluye tanto a individuos con un grado de VISIÓN ESTÁNDAR como a pacientes con BAJA VISIÓN.

ABSTRACT

This Ph. Thesis aims to provide an efficient natural scene feature extraction technique for Low VISION patients. This work has been developed with this motivation in mind. It provides two main scientific contributions. The first one is the adaptation of the *convolution* operator to the LIP paradigm. It allows to adapt traditional methods based on separable filters to the LIP paradigm. Due to the logarithmic behaviour of LIP, boundaries extraction algorithms highly invariant to illumination changes can be designed. Throughout this Ph. Thesis, several traditional edge extraction algorithms are adapted to LIP paradigm by using the new proposed operator.

The second contribution is the SUBJECTIVE evaluation of the perceived quality of the contour maps obtained by each method using questionnaires. In this Ph. Thesis, several recommendations for the design of good contour images evaluation surveys have been proposed. By using these recommendations, a survey has been designed to evaluate the quality of the methods adapted to LIP using the new proposed operator. The population sample that answers that survey includes both STANDARD LEVEL VISION persons and Low VISION patients.

PUBLICACIONES CIENTÍFICAS

Algunas de las ideas, imágenes y datos que se exponen en esta Tesis Doctoral han aparecido publicadas previamente en algunas publicaciones:

- **Rafael Palomar, José M. Palomares, José M. Castillo, Joaquín Olivares y Juan Gómez-Luna** *Parallelizing and Optimizing LIP-Canny Using NVIDIA CUDA* En N. García-Pedrajas et al. (editor): *Proceedings of the IEA/AIE 2010, Córdoba (España)* publicado en *Lecture Notes in Artificial Intelligence*, LNAI 6098(III): 389–398, 2010.
- **José M. Palomares, Jesús González, Eduardo Ros y Alberto Prieto.** *General Logarithmic Image Processing Convolution*. *IEEE Transactions on Image Processing*, 15(11):3602–3608, 2006. DOI: [10.1109/tip.2006.881967](https://doi.org/10.1109/tip.2006.881967). Índice de Impacto: 2.715
- **José M. Palomares, Jesús González y Eduardo Ros.** *Designing a fast convolution under the LIP paradigm applied to edge detection* En Sameer Singh et al. (editor): *Proceedings of the III ICAPR 2005, Bath (Reino Unido)* publicado en *Lecture Notes on Computer Sciences*, LNCS 3687(3):560–569, 2005. DOI: [10.1007/11552499_62](https://doi.org/10.1007/11552499_62). Índice de Impacto: 0.402
- **José M. Palomares, Jesús González y Eduardo Ros.** *Detección de bordes en imágenes con sombras mediante LIP-Canny* En las Actas del I Simposio de Reconocimiento de Formas y Análisis de Imágenes: *Conferencia AERFAI 2005, Granada (España)*:71–76, 2005. ISBN:84-9732-445-5

*Todas las verdades son fáciles de entender, una vez descubiertas.
El caso es descubrirlas.*
— Galileo Galilei

AGRADECIMIENTOS

LISTA DE CONTENIDOS

I	INTRODUCCIÓN	1
1	INTRODUCCIÓN	3
1.1	Motivación	4
1.1.1	Procesamiento de imágenes	5
1.1.2	Iluminación no uniforme	6
1.1.3	La calidad en el procesamiento de imágenes	6
1.2	Campos de aplicación	7
1.2.1	Sistemas de ayuda para Baja Visión	7
1.2.2	Sistemas inteligentes de Realidad Aumentada	11
1.3	Marco de trabajo	14
1.3.1	Calidad de las imágenes procesadas	14
1.3.2	Imágenes en Tiempo Real	14
1.4	Planteamiento	15
1.4.1	Problemática	15
1.4.2	Ámbito de actuación	15
1.4.3	Objetivos científicos de la Tesis	15
1.4.4	Propuesta inicial	16
1.5	Estructura de la Tesis	17
2	TÉCNICAS DE PROCESAMIENTO DE IMÁGENES	21
2.1	Modelos de formación de imágenes	22
2.1.1	Filtrado homomórfico	22
2.2	Descripción del modelo LIP	24
2.2.1	Nomenclaturas y rangos	25
2.2.2	Bases matemáticas de LIP	26
2.2.3	Relación con otros modelos logarítmicos	29
2.2.4	Leyes físicas y psico-físicas que cumple LIP	30
2.3	Modelos de ruido en imágenes	33
2.3.1	Ruido en imágenes	33
2.3.2	Emborronamiento de imágenes	34
2.4	Extracción de bordes	35
2.4.1	Filtros basados en el gradiente de la imagen	36
2.4.2	Filtros basados en el Laplaciano de la imagen	38
2.4.3	Método de Canny para la extracción de bordes	41
2.5	Conclusiones	43
3	EVALUACIÓN DE LA CALIDAD DE IMÁGENES PROCESADAS	45
3.1	Evaluación de resultados	46
3.1.1	Tipos de evaluación de resultados	47
3.1.2	Evaluación objetiva	48
3.1.3	Evaluación subjetiva	49
3.1.4	Restricciones generales	50
3.2	Contornos de los objetos en imágenes	51
3.2.1	Importancia de los contornos en imágenes	51
3.2.2	Conceptos de los contornos en imágenes	51
3.3	Evaluación de la calidad de contornos	52
3.3.1	Evaluación objetiva de los contornos de imágenes	52

3.3.2	Evaluación subjetiva de los contornos de imágenes	54	
3.4	Mean Opinion Score: MOS	56	
3.4.1	Cálculo del MOS	58	
3.4.2	Extensión del MOS	58	
3.5	Metodología de encuestas de imágenes de contornos		60
3.6	Conclusiones	62	
4	ENCUESTAS	63	
4.1	Encuestas de opinión mediante cuestionarios		64
4.2	¿Qué es un cuestionario?	65	
4.2.1	Definición de qué es un cuestionario	65	
4.2.2	Características buscadas con los cuestionarios		67
4.3	Mecanismos de recolección de las respuestas		67
4.3.1	Entrevistas estructuradas	67	
4.3.2	Cuestionarios auto-administrados	68	
4.3.3	Tipos de preguntas	69	
4.4	Limitaciones de los cuestionarios	70	
4.5	Fuentes de error en los cuestionarios	71	
4.6	Elección de la muestra poblacional	74	
4.6.1	Selección de individuos	74	
4.7	Diseño de un cuestionario	76	
4.7.1	Dificultades para construir buenas preguntas		76
4.7.2	Buenas preguntas con buenas respuestas		79
4.8	Análisis del cuestionario	82	
4.8.1	Validación del cuestionario	82	
4.8.2	Fiabilidad del cuestionario	83	
4.8.3	Alfa de Cronbach	83	
4.8.4	Coefficiente de Correlación Intraclase		85
4.9	Conclusiones	87	
5	CUESTIONARIOS SOBRE BAJA VISIÓN	89	
5.1	Investigación sobre la BAJA VISIÓN mediante cuestionarios	90	
5.1.1	Visual Function Index	90	
5.1.2	Rand Questionnaire to Assess Functional Problems of the Visually Impaired		91
5.1.3	Visual Function after Pan-Retinal Photocoagulation		91
5.1.4	Visual Status Inventory	92	
5.1.5	Lowe's Visual Function	92	
5.1.6	Questionnaire for Functional Assessment of Low Vision		92
5.1.7	Visual Activities Questionnaire	93	
5.1.8	Visual Performance Questionnaire	93	
5.1.9	Activities of Daily Vision Scale	94	
5.1.10	Assessment of Visual Function-Related Quality of Life		94
5.1.11	14-Item Visual Functioning Index	95	
5.1.12	Vision-Related Quality of Life Questionnaire		95
5.1.13	National Eye Institute's Visual Functioning Questionnaire		96
5.1.14	Cuestionario de Calidad Visual del Consejo Argentino de Oftalmología		97

5.2	Dominios relevantes para BAJA VISIÓN y procesamiento de imágenes	98
5.3	Conclusiones	100
5.3.1	Conclusiones de la Parte I	100
II MÉTODO Y MATERIALES 101		
6	CONVOLUCIÓN LIP	103
6.1	La convolución en el procesamiento de imágenes	104
6.1.1	Descripción matemática de la convolución	104
6.1.2	Convolución 2D con filtros separables	105
6.2	Convolución LIP 2D con filtros separables	106
6.2.1	Desarrollo matemático de la Convolución LIP 2D	106
6.2.2	Convolución LIP Generalizada Directa: DGLIP-Conv	108
6.2.3	Convolución LIP Generalizada Rápida: FGLIP-Conv	109
6.3	Rango de funcionamiento de la Convolución LIP	109
6.3.1	Límite inferior del rango	109
6.3.2	Límite superior del rango	110
6.3.3	Límite inferior y superior del rango	111
6.4	Comparativas de rendimiento	111
6.4.1	Coste computacional por operador	113
6.5	Conclusiones	114
7	CRITERIOS PARA LA EVALUACIÓN DE LA CALIDAD DE CONTORNOS	117
7.1	Objetivo principal y Criterios generales	118
7.1.1	Criterios generales	118
7.2	Criterios específicos	118
7.2.1	DOMINIOS del cuestionario	119
7.2.2	Contexto de las tareas de visión	120
7.2.3	Selección de las técnicas de detección de bordes	120
7.2.4	Selección del ámbito de actuación	122
7.2.5	Selección de los parámetros visuales	123
7.3	Imágenes para el cuestionario de evaluación en BAJA VISIÓN	124
7.4	Conclusiones	134
7.4.1	Contexto del capítulo	134
8	RECOMENDACIONES PARA EL DISEÑO DE CUESTIONARIOS DE CALIDAD DE CONTORNOS	137
8.1	Sesgo por finalidad	138
8.2	Recomendaciones generales sobre las preguntas	139
8.2.1	Estructura de las preguntas	139
8.2.2	Imágenes respuesta	140
8.2.3	Ponderación de preguntas	140
8.3	Tipos de preguntas con elementos visuales	141
8.3.1	Preguntas de SELECCIÓN O DICOTÓMICAS	141
8.3.2	Preguntas de PUNTUACIÓN	142
8.3.3	Preguntas de ORDENAMIENTO	143
8.3.4	Preguntas de MAPEADO ESPACIAL Y TEMPORAL	143
8.4	Recomendaciones sobre la estructura general del cuestionario	145

8.4.1	Reutilización de imágenes	145
8.4.2	Uso de imágenes reales	146
8.4.3	Número de <i>imágenes respuesta</i> por pregunta	147
8.5	Sistema Web para cuestionarios multimedia	148
8.5.1	Preguntas de PUNTUACIÓN	150
8.5.2	Preguntas de SELECCIÓN SIN REPETICIÓN	150
8.5.3	Preguntas de SELECCIÓN CON REPETICIÓN	151
8.5.4	Preguntas de MAPEADO ESPACIAL Y TEMPORAL	151
8.6	Conclusiones	152
8.6.1	Conclusiones de la Parte II	152
III	RESULTADOS EXPERIMENTALES	155
9	APLICACIONES CONVOLUCIÓN LIP	157
9.1	LIP-Sobel	158
9.1.1	Formulación	158
9.1.2	Resultados experimentales	159
9.2	Filtros de emborronamiento	161
9.2.1	Filtro Media: LIP-Avg	162
9.2.2	Filtro Gaussiano: LIP-GBLur	163
9.3	Filtros para la extracción de bordes	165
9.3.1	Laplaciano de Gaussianas: LIP-LoG	165
9.3.2	LIP-Canny	169
9.4	Tiempos de cómputo	171
9.4.1	Aumento del rendimiento (<i>Speedup</i>)	172
9.4.2	LIP-Sobel	172
9.4.3	Algoritmos de emborronamiento: LIP-Avg, LIP-GBLur	173
9.4.4	Extracción de bordes: LIP-LoG	174
9.4.5	Extracción de bordes: LIP-Canny	175
9.5	Resumen de resultados	176
9.5.1	LIP-Sobel	177
9.5.2	LIP-Avg, LIP-GBLur	178
9.5.3	LIP-LoG	179
9.5.4	LIP-Canny	180
9.6	Conclusiones	180
10	RESULTADOS DE LOS CUESTIONARIOS DE CALIDAD DE BORDES	183
10.1	Cuestionarios para evaluación de contornos	184
10.2	Pilotaje	185
10.2.1	Pilotaje previo	185
10.2.2	Pilotaje final	186
10.3	Visibilidad de la encuesta	187
10.4	Encuestados	188
10.4.1	Muestra poblacional	188
10.4.2	Carta de presentación	189
10.4.3	Tasa de respuesta	192
10.5	Resultados del proceso de evaluación	194
10.5.1	Respuestas cualitativas	194
10.5.2	Metodología de las respuestas	195
10.6	Análisis preliminar de los cuestionarios	196
10.6.1	Análisis de la población	196
10.6.2	Muestreo a posteriori	197
10.6.3	Análisis de frecuencia de cada pregunta	198

10.7	Análisis de los resultados	200
10.7.1	Análisis por grado de visión de los encuestados	200
10.8	Validación de la encuesta	205
10.8.1	Validez de criterio	206
10.8.2	Validez de constructo	206
10.8.3	Validez de contenido	207
10.9	Fiabilidad de la encuesta	208
10.9.1	Fiabilidad externa	208
10.9.2	Fiabilidad interna	208
10.10	Conclusiones	211
10.10.1	Conclusiones sobre el proceso de respuesta	212
10.10.2	Conclusiones sobre los resultados	214
10.10.3	Conclusiones de la Parte III	216
IV DISCUSIÓN Y CONCLUSIONES 217		
11	DISCUSIÓN	219
11.1	Discusión sobre el operador <i>Convolución</i> -LIP	220
11.1.1	Doble implementación	220
11.1.2	Fácil adaptación a LIP	221
11.1.3	Mejor comportamiento en zonas con baja iluminación	221
11.1.4	Ajuste robusto de parámetros	221
11.1.5	Comportamiento unificado sin pasos adicionales	222
11.2	Discusión sobre las encuestas de evaluación de contornos	222
11.2.1	Métodos LIP mejor valorados	223
11.2.2	Métodos LIP aprueban	223
11.2.3	Encuestados puntúan de manera similar	224
11.2.4	Respuestas proporcionales entre grupos	224
11.3	Conclusiones	224
12	CONCLUSIONES	227
12.1	Sumario	228
12.2	Producción científica	232
12.3	Principales aportaciones científicas	233
12.3.1	Sobre el operador <i>Convolución</i> -LIP	233
12.3.2	Sobre las encuestas de evaluación de contornos	233
12.4	Investigaciones futuras	235
V APÉNDICES 237		
A	APÉNDICE	239
A.1	Demostración del rango de funcionamiento de la <i>Suma</i> -LIP	239
A.1.1	Valores superiores a cero	240
A.1.2	Valores inferiores a M	240
A.1.3	Rango completo	241
B	APÉNDICE	243
B.1	Demostración del rango de funcionamiento de la <i>Multiplicación</i> -LIP	243
B.1.1	Valores superiores a cero	244
B.1.2	Valores inferiores a M	244
B.1.3	Rango completo	245

C	APÉNDICE	247
c.1	Cuestionarios utilizados en el estudio	248
c.1.1	Cuestionario 1	248
c.1.2	Cuestionario 2	254
	BIBLIOGRAFÍA	261

LISTA DE FIGURAS

Figura 2.1	Imágenes de ejemplo: <i>Muestra las dos imágenes de ejemplo utilizadas para mostrar el funcionamiento de LIP.</i>	27
Figura 2.2	Suma de dos imágenes: <i>Comparación de la operación suma de dos imágenes entre la suma estándar y la suma LIP.</i>	27
Figura 2.3	Multiplicación escalar por 4: <i>Comparación de la operación multiplicación escalar de una imagen por el valor 4, utilizando la multiplicación estándar y la multiplicación LIP.</i>	27
Figura 3.1	Calidad ACR 11 niveles con anclaje: <i>Escala gráfica para la calificación absoluta de la calidad con puntos de anclaje para 11 niveles de calidad.</i>	57
Figura 3.2	ITEM del cuestionario de HEATH <i>et al.</i> [49]: <i>Imagen-pregunta para la evaluación de la calidad de bordes.</i>	61
Figura 6.1	Comparativa FGLIP-CONV vs. DGLIP-CONV	113
Figura 6.2	Speedup FGLIP-CONV vs. DGLIP-CONV	113
Figura 7.1	Imagen Personas 1	125
Figura 7.2	Imagen Personas 2	125
Figura 7.3	Imagen Personas 3	125
Figura 7.4	Imagen Personas 4	126
Figura 7.5	Imagen Personas 5	126
Figura 7.6	Imagen Cercanas 1	127
Figura 7.7	Imagen Cercanas 2	127
Figura 7.8	Imagen Cercanas 3	128
Figura 7.9	Imagen Cercanas 4	128
Figura 7.10	Imagen Cercanas 5	129
Figura 7.11	Imagen Interiores 1	129
Figura 7.12	Imagen Interiores 2	130
Figura 7.13	Imagen Interiores 3	130
Figura 7.14	Imagen Interiores 4	131
Figura 7.15	Imagen Interiores 5	131
Figura 7.16	Imagen Exteriores 1	132
Figura 7.17	Imagen Exteriores 2	132
Figura 7.18	Imagen Exteriores 3	132
Figura 7.19	Imagen Exteriores 4	133
Figura 7.20	Imagen Exteriores 5	133
Figura 8.1	Encuestas trabajo Heath	149
Figura 8.2	Pregunta de PUNTUACIÓN	150
Figura 8.3	Pregunta de SELECCIÓN SIN REPETICIÓN	150
Figura 8.4	Pregunta de SELECCIÓN CON REPETICIÓN	151
Figura 8.5	Pregunta de MAPPING	151
Figura 9.1	Imagen Original	159
Figura 9.2	Imagen Oscurecida	160
Figura 9.3	Imagen Sobel	160
Figura 9.4	Imagen LIP-Sobel	160
Figura 9.5	Lenna Original	167

Figura 9.6	Lenna Oscurecida	167
Figura 9.7	LoG Lenna Oscurecida	167
Figura 9.8	LoG umbralizado Lenna Oscurecida	168
Figura 9.9	LOG–LIP Lenna Oscurecida	168
Figura 9.10	LOG–LIP umbralizado Lenna Oscurecida	168
Figura 9.11	Sombras Original	169
Figura 9.12	Canny	170
Figura 9.13	<i>Filtrado Homomórfico</i> +Canny	170
Figura 9.14	LIP–Canny	170
Figura 10.1	Gráfica respuesta encuestas	193
Figura 10.2	Gráfica respuesta métodos	201
Figura 10.3	Gráfica respuesta métodos	202
Figura 10.4	Gráfica regresión lineal según tipo de visión	205
Figura 10.5	Gráfica de valoración media de cada usuario por método	211

LISTA DE TABLAS

Tabla 1.1	Patologías y Baja Visión	9
Tabla 3.1	Calidad ACR 5 niveles	57
Tabla 3.2	Calidad ACR 9 niveles	57
Tabla 3.3	Criterios generales calidad	60
Tabla 4.1	Valoración cualitativa del ICC	87
Tabla 5.1	Determinación de DOMINIOS para BAJA VISIÓN	99
Tabla 6.1	Coste computacional por operador.	114
Tabla 7.1	Objetivo del cuestionario	118
Tabla 7.2	Dominios seleccionados	120
Tabla 7.3	Tareas típicas de visión	121
Tabla 7.4	Técnicas seleccionadas	122
Tabla 7.5	Objetivos principales de la encuesta	122
Tabla 7.6	Grupos poblacionales de la encuesta	123
Tabla 7.7	Características homogéneas de todos los métodos evaluados en la encuesta.	123
Tabla 7.8	Determinación de umbrales	124
Tabla 9.1	Gauss 7×7	164
Tabla 9.2	Tiempos cálculo	172
Tabla 9.3	Speedup LIP–Sobel	173
Tabla 9.4	Tiempos cálculo	174
Tabla 9.5	Speedup LIP–Avg, LIP–GBlur	174
Tabla 9.6	Tiempos LoG y LIP–LoG	175
Tabla 9.7	Speedup LoG y LIP–LoG	175
Tabla 9.8	Tiempos Canny, LIP–Canny y FH+Canny	176
Tabla 9.9	Speedup Canny, LIP–Canny y FH+Canny	176
Tabla 10.1	Respuestas encuesta	193
Tabla 10.2	Tasa Respuesta de la Muestra Poblacional Estimada	194
Tabla 10.3	Detección de outliers	199

Tabla 10.4	Puntuación subjetiva de cada método	201
Tabla 10.5	Puntuación subjetiva según grado de visión	202
Tabla 10.6	Tabla de correlación entre grados de visión	203
Tabla 10.7	Puntuación subjetiva según grado de visión	204
Tabla 10.8	Tabla de correlación entre grados de visión	204
Tabla 10.9	Puntuación de validez de contenido	207
Tabla 10.10	α -Cronbach de los Cuestionarios	209
Tabla 10.11	α -Cronbach de los Métodos según preguntas	210
Tabla 10.12	ICC de los Cuestionarios según tipo de usuario	211
Tabla 11.1	Características <i>Convolución</i> -LIP	220
Tabla 11.2	Características evaluación de contornos por encuestas	223

ACRÓNIMOS

ADVS	Activities of Daily Vision Scale
CAO	Consejo Argentino de Oftalmología
CCD	Charge-Coupled Device
CCV-07-MUDES	Cuestionario de Calidad Visual-07 con capacidad residual Menor de Una Décima en la Escala de Snellen
DGLIP-Conv	Direct General LIP-Convolution
FALV	Questionnaire for Functional Assessment of Low Vision
FGLIP-Conv	Fast General LIP-Convolution
GANIP	General Adaptive Neighborhood Image Processing
ICC	Intraclass Correlation Coefficient
LIP	Logarithmic Image Processing
LoG	Laplacian of Gaussian
LRIP	Log-ratio Image Processing
LVF	Lowe's Visual Function
MHIP	Multiplicative Homomorphic Image Processing
MOS	Mean Opinion Score
MSE	Mean Squared Error
NEI	National Eye Institute
NEI-VFQ	National Eye Institute's Visual Functioning Questionnaire
OMS	Organización Mundial de la Salud

ONCE	Organización Nacional de Ciegos Españoles
PCA	Principal Component Analysis
PIXEL	PI(X)cture ELement
PSNR	Peak Signal-to-Noise Ratio
RAE	Real Academia Española de la lengua
Rand-FPVI	Rand Questionnaire to assess Functional Problems of the Visually Impaired
UIT	Unión Internacional de Telecomunicaciones
UIT-T	Unión Internacional de Telecomunicaciones, sección normalización de Telecomunicaciones
VAQ	Visual Activities Questionnaire
VCM ₁	Vision quality-of-life Core Measure 1
VFI	Visual Function Index
VF-PRP	Visual Function after Pan-Retinal Photocoagulation
VF-QOL	Visual Function-related Quality of Life
VF-14	Visual Functioning index with 14-item
VPQ	Visual Performance Questionnaire
VSI	Visual Status Inventory

Parte I

INTRODUCCIÓN

INTRODUCCIÓN

ÍNDICE

1.1	Motivación	4	
1.1.1	Procesamiento de imágenes	5	
1.1.2	Iluminación no uniforme	6	
1.1.3	La calidad en el procesamiento de imágenes	6	
1.2	Campos de aplicación	7	
1.2.1	Sistemas de ayuda para Baja Visión	7	
1.2.2	Sistemas inteligentes de Realidad Aumentada	11	
1.3	Marco de trabajo	14	
1.3.1	Calidad de las imágenes procesadas	14	
1.3.2	Imágenes en Tiempo Real	14	
1.4	Planteamiento	15	
1.4.1	Problemática	15	
1.4.2	Ámbito de actuación	15	
1.4.3	Objetivos científicos de la Tesis	15	
1.4.4	Propuesta inicial	16	
1.5	Estructura de la Tesis	17	

La evolución de la Ciencia y de la Tecnología proviene de las contribuciones de muchos investigadores a lo largo de la Historia. Los trabajos de todos estos investigadores se han desarrollado siguiendo principalmente dos tipos de investigaciones: la investigación teórica y la aplicada. Para el desarrollo de la Ciencia y de la sociedad, ambas son importantes y además complementarias entre sí, no excluyendo la una a la otra. En esta Tesis en particular, se ha querido desarrollar principalmente un trabajo de investigación aplicada, aunque sin dejar de lado el aspecto teórico. Para conseguir este fin se ha partido de una primera motivación de tipo social en la que un determinado colectivo plantea una problemática concreta. Diversos investigadores han propuesto algunas soluciones para resolver esta problemática, que sin embargo, presentan ciertas limitaciones. La búsqueda de soluciones a estas limitaciones ha sido la motivación técnica que ha alentado esta Tesis Doctoral. Así pues, esta Tesis ha pretendido dar solución a un problema científico-técnico concreto enmarcado dentro de un contexto de aplicación social más general.

Esta Tesis Doctoral involucra tanto Investigación Aplicada como Investigación Teórica.

La memoria de la presente Tesis Doctoral ha sido dividida en diferentes partes, que a su vez están subdivididas en capítulos. Esta primera parte de la Tesis Doctoral sirve de introducción del trabajo, que se expone y desarrolla en el resto de partes y capítulos. Esta parte se inicia describiendo una problemática presente en la sociedad, los impedimentos que sufren los pacientes con problemas de BAJA VISIÓN en su vida diaria.

Estos pacientes, que presentan una agudeza visual reducida y/o un campo de visión mucho menor del habitual, sufren grandes dificultades para orientarse y en muchos casos no pueden valerse por sí mismos. Para ayudar a estos pacientes se han construido gafas y otros mecanismos [5, 83, 84], tanto ópticos como electrónicos, que permiten ampliar su campo de visión o aumentar la definición de los objetos que observan. Sin embargo, los sistemas de tipo óptico requieren una adaptación y entrenamiento por parte de los pacientes para obtener un grado de mejora bastante limitado. Por otra parte, con los sistemas de tipo electrónico se obtienen mejores resultados y se consigue una elevada aceptación de los mismos por parte de los pacientes, pero presentan varios inconvenientes: son voluminosos y por el momento, no están diseñados para ser utilizados fuera del ámbito doméstico o en entornos muy controlados, principalmente la lectura de libros [5] o para hacer *zoom*, tanto de imágenes de vídeo como reales [21]. Algunos científicos han construido prototipos [94, 113] que hacen uso de cámaras digitales y algoritmos de procesamiento de imágenes para ayuda a la navegación en pacientes con BAJA VISIÓN, con buenos resultados en general. No obstante presentan algunos problemas, entre los que destacan la detección de objetos en zonas con sombras proyectadas o en zonas con iluminación no uniforme. La búsqueda de una solución plausible que resuelva todos estos problemas se describe a lo largo de la Tesis y ha sido la motivación de este trabajo. En esta primera parte, también se explica cómo esta solución podría tener aplicación en otros campos diferentes. Tomando como campo de aplicación el de la ayuda a los pacientes con BAJA VISIÓN, se ha acotado el marco de trabajo, lo que a su vez, ha posibilitado una implementación concreta.

1.1 MOTIVACIÓN SOCIAL Y PROBLEMÁTICA CIENTÍFICO-TÉCNICA

Teniendo en cuenta que en esta Tesis Doctoral se exponen técnicas de procesamiento digital de imágenes para resolver problemas de visión, es conveniente realizar una breve exposición de este ámbito de trabajo e investigación. Existen muchas aplicaciones de visión artificial en la actualidad que se utilizan en diversos entornos: cámaras de control de tráfico [34], sensores de presencia [73, 111], mecanismos autónomos de guía de vehículos [27], retinas artificiales [57], etc. La disciplina de “visión artificial” es un campo de investigación interdisciplinar en constante avance en el que muchos investigadores de distintas ramas del saber se están dedicando de manera directa o indirecta a obtener nuevos métodos o a mejorar los mecanismos de procesamiento de los mismos. Estas mejoras en los resultados obtenidos permiten construir aplicaciones de mayor complejidad.

1.1.1 *Visión artificial y procesamiento de imágenes y vídeo*

Existen multitud de esquemas de visión artificial estructurados de diferentes formas. Uno de ellos se puede describir mediante un esquema simplificado, en analogía con el funcionamiento de los compiladores. Los algoritmos de visión artificial suelen comenzar con una fase, denominada ANÁLISIS DE LA ESCENA, que suele, a su vez, estar subdividida en varias etapas consecutivas. En la primera etapa de esta fase, se realiza un procesamiento de bajo nivel de la imagen (el equivalente al procesamiento de nivel léxico de los compiladores), gracias al cual se extraen los elementos estructurales más básicos y sencillos, que constituyen los componentes fundamentales de las imágenes captadas. Una segunda etapa de esta fase, conocida como RECONOCIMIENTO DE PATRONES, busca interrelaciones y agrupaciones complejas entre los diversos elementos básicos, extraídos en la etapa anterior (esta etapa equivaldría al nivel sintáctico de los compiladores). En el último nivel de análisis, que se suele denominar DETECCIÓN O IDENTIFICACIÓN DE OBJETOS, se estudian dichas agrupaciones para extraer las estructuras coherentes con significado en el mundo real en las que se puede integrar información de modelos del mundo (aspecto similar al nivel semántico de los compiladores). Estos modelos se han generado a partir de otras experiencias anteriores o de bases de datos bien estructuradas. Los algoritmos inteligentes de toma de decisiones que se implementan en niveles superiores utilizan los resultados obtenidos en esta fase de análisis.

Centrándose en la primera etapa, de EXTRACCIÓN DE CARACTERÍSTICAS, existen muchas técnicas que nos permiten extraer algunas propiedades fundamentales de las imágenes para poder caracterizarlas. De entre todas las técnicas existentes, las que más se suelen aplicar son las de DETECCIÓN DE BORDES O CONTORNOS DE OBJETOS. Existen muchas implementaciones distintas para la detección del borde de los objetos, cada una de las cuales está basada en técnicas diferentes. Cabe preguntarse por qué, de entre todos los métodos posibles, los algoritmos de detección de bordes se suelen aplicar en las primeras fases de la extracción de características: está ampliamente documentado [108–110] que el ojo humano y las primeras etapas de procesamiento visual humano van equipadas con estructuras neuronales que extraen de forma eficiente zonas de alto contraste, que habitualmente coinciden con los bordes de los objetos. Es decir, el sistema visual humano utiliza filtros espacio-temporales para esas primeras etapas de procesamiento. De esta manera, aplicando algoritmos de extracción de bordes podemos reducir el volumen de datos que hay que procesar, puesto que todas las zonas interiores de los objetos no contienen información de bordes y se pueden descartar. Por tanto, las distintas técnicas de detección de contornos nos van a extraer de forma eficiente y rápida estas características espaciales (de alto contraste local) como zonas con mayor información, que serán utilizadas por los algoritmos de los niveles superiores. Estas técnicas también permiten eliminar muchas zonas espacialmente redundantes y con *relativamente* poca información de interés para el sistema visual humano.

Las diferentes técnicas e implementaciones de los detectores de contornos de los objetos suelen representar el resultado final mediante un

La detección de bordes de objetos es la técnica más utilizada en la etapa de extracción de características

PIXEL: Del inglés, PI(X)cture ELeMent; se suele traducir como PUNTO DE LA IMAGEN

mapa binario de igual tamaño que la imagen original, en el que el valor binario de cada PIXEL de la imagen indica su pertenencia a los bordes de alguno de los objetos presentes en la escena. El referido mapa binario suele ser el refinamiento final proveniente de la umbralización de uno o más mapas previos que contienen diversas “medidas” (por ejemplo: magnitud del gradiente, laplaciano del gradiente, entropía, etc.). La umbralización puede ser tan simple como la comparación de cada PÍXEL de la imagen con un valor único, o puede ser mucho más compleja y usar valores de umbral localmente adaptativos y ventanas de tamaño dinámico. El uso de umbrales globales acelera los cálculos a costa de tener resultados muy sensibles al ruido, mientras que el uso de umbrales más complejos provee resultados más robustos frente al ruido, aunque mucho más costosos en tiempo de procesamiento.

1.1.2 La problemática de la iluminación no uniforme

El problema de la iluminación no uniforme de las escenas provoca degradación de los resultados en el procesamiento de las escenas.

Prácticamente desde los inicios de las disciplinas de visión artificial y de procesamiento de imágenes y vídeo, diversos autores [17, 77, 78] han informado del problema de la iluminación no uniforme de las escenas, que suele llevar a un deficiente procesamiento de los elementos presentes en las mismas. Las sombras proyectadas o los focos de luz pueden interferir en los algoritmos alterando el resultado de manera “artificiosa”. Esta problemática se observa muy claramente en la detección de los bordes de los objetos en zonas de sombras proyectadas. En estos casos es necesario recurrir a algoritmos muy complejos o a técnicas de umbralización adaptativas, que hacen muy complicado el proceso de selección de los parámetros que optimizan el resultado. Así pues, la iluminación no uniforme, las sombras proyectadas, los focos de luz y elementos similares que afecten a la iluminación incidente en una determinada zona de la imagen provoca, en la mayoría de los casos, la no detección en absoluto de fronteras en dichas zonas o, en el mejor de los casos, una detección muy deteriorada de las mismas. Esto suele contrastar con las fronteras detectadas en las zonas bien iluminadas, que tienen valores mucho más elevados, lo que suele provocar que los bajos valores de las fronteras de las zonas de baja iluminación se cataloguen como “falsos positivos” y se descarten del mapa de bordes final. Por tanto, todos aquellos mecanismos que proporcionen un resultado invariante ante los cambios de iluminación natural, proporcionarán una mayor calidad en la detección de bordes de imágenes.

1.1.3 La calidad en el procesamiento de imágenes

Al contrario que en muchas de las áreas de investigación de las Ingenierías, para el PROCESAMIENTO DE IMÁGENES Y VÍDEO en general, y para la VISIÓN ARTIFICIAL en particular, no se suele conocer *a priori* el resultado “perfecto”. Esto es así porque la bondad de un resultado depende en gran medida del objetivo que se persiga. Por ejemplo, para una imagen de una carretera captada por un robot de ayuda de navegación, basta con que se detecten las líneas rectas que delimitan los carriles de la carretera, aunque se pierdan todos los matices y detalles como árboles, casas, paisajes, etc.; sin embargo, para un sistema de detección de caras lo importante puede ser que detecte perfectamente los

La calidad de las imágenes procesadas se evalúa de manera indirecta

rasgos faciales, aunque no sea capaz de detectar ningún otro elemento del cuerpo. Es por ello que la medición de la calidad de las imágenes procesadas se suele hacer de manera indirecta. Para ello, se mide la bondad de la respuesta de un mismo método de niveles superiores cambiando los métodos de PROCESAMIENTO DE IMÁGENES Y VÍDEO de los niveles inferiores. Se da por supuesto que si el mismo módulo es capaz de mejorar el resultado obtenido, es porque el método que ha extraído las características tiene un mejor comportamiento.

Sin embargo, existen casos en los que no es posible hacer ninguna medición indirecta. Por ejemplo, cuando se prueban de manera independiente métodos de PROCESAMIENTO DE IMÁGENES Y VÍDEO en los que *a priori* no se conoce para qué se van a utilizar; bien porque aún no se dispongan de los módulos de procesamiento de niveles superiores, o bien porque el usuario final va a visualizar directamente las imágenes procesadas por estos métodos sin ningún procesamiento digital posterior. En todos estos casos, el usuario es el único “juez” válido para medir la calidad de las imágenes procesadas, y además, la respuesta que proporcione se basará únicamente en su opinión, ya que no podrá adjuntar ninguna medición cuantitativa. Para la obtención de la opinión de los usuarios, el mecanismo más utilizado consiste en el análisis de encuestas de opinión.

Existen algunos protocolos o estándares para la cumplimentación de encuestas de opinión que contengan imágenes o vídeo para su evaluación cualitativa [53]. Sin embargo, estos estándares sólo proporcionan información de cómo se deben presentar las imágenes o los vídeos que se desean evaluar (posición del observador, iluminación de la estancia, etc.) y no presentan recomendaciones con respecto a la selección de la muestra poblacional ni a la construcción y desarrollo de las encuestas. Esto puede provocar que las encuestas estén muy sesgadas, por lo que el resultado podría no ser generalizable y, por tanto, las conclusiones obtenidas no serían válidas.

Cuando la calidad de las imágenes procesadas no es medible indirectamente, se realizan encuestas subjetivas

1.2 CAMPOS DE APLICACIÓN

Como se ha indicado previamente, la motivación principal del trabajo realizado en esta Tesis Doctoral es propiciar una aplicación real de la investigación científico-técnica que sea útil a los intereses de la sociedad. Por ello, a continuación se indican algunos campos de aplicación en los que la inclusión de las técnicas objeto de estudio en esta Tesis podrían significar un claro avance. Tras un análisis de estos campos de aplicación se ha delimitado el marco de trabajo, que ha servido para obtener dos tipos de características, las cuales acotan y determinan el trabajo de investigación: los objetivos científicos que deben cumplir los métodos diseñados y las restricciones técnicas que la tecnología actual impone en la implementación de los algoritmos desarrollados.

1.2.1 *Sistemas de ayuda para personas con patologías visuales: BAJA VISIÓN*

Existen muchos grados y causas por las que una persona puede sufrir algún tipo de deficiencia visual. La graduación puede ir desde una

degradación de la calidad de visión en general (presbicia) o en los objetos situados a corta o a larga distancia (hipermetropía o miopía, respectivamente), hasta la ceguera total o parcial de uno o de ambos ojos, pasando por diversas deformaciones y aberraciones ópticas de los objetos (como por ejemplo pasa con el astigmatismo). Además, estos problemas pueden estar causados por defectos anatómicos o por lesiones en diferentes partes del sistema visual, como:

- A. El globo ocular (principalmente, en la córnea y el cristalino).
- B. El fondo del ojo (la retina).
- C. El nervio óptico.
- D. El **CÓRTEX VISUAL**.

CÓRTEX VISUAL: La parte del cerebro que analiza e interpreta las imágenes.

Los enfermos que sufren alguna clase de degeneración del sistema visual suelen presentar problemas, en mayor o menor medida, para poder manejarse por sí mismos, debido a que el campo visual ha sufrido una gran disminución (en algunos casos, completamente) o a que la percepción de los escenarios se ha empobrecido hasta tal punto que no son capaces de determinar la presencia de ningún elemento en particular.

BAJA VISIÓN *“Se considera que una persona tiene BAJA VISIÓN cuando padece una limitación visual que le dificulta o impide la realización de las tareas de la vida cotidiana. La BAJA VISIÓN puede ser causada por distintas patologías, accidentes o malformaciones congénitas. En todos los casos se trata de una razón crónica e irreversible, por lo que la rehabilitación de la BAJA VISIÓN entra en juego cuando el oftalmólogo no puede hacer nada más por mejorar la visión.”* [6].

OMS: Organización Mundial de la Salud

De manera más precisa científicamente, la *Organización Mundial de la Salud*, **OMS**, definió en 1992 que se considera que un paciente sufre **BAJA VISIÓN** si en el mejor ojo con la mejor corrección posible su *agudeza visual* desciende por debajo de 6/20 (o 0.3) en la escala de Snellen o tiene un *campo visual* menor de 20° [5, 75].

La *agudeza visual* mide la capacidad para percibir la figura y la forma de los objetos así como para discriminar sus detalles. Para medirla se utilizan generalmente optotipos o paneles de letras o símbolos, para lo que habitualmente se utiliza el test de Snellen (ver **TEST DE SNELLEN** en el pie de la página 9). El *campo visual* se define como la capacidad para percibir los objetos situados fuera de la visión central (que corresponde al punto de visión más nítido). La valoración del campo visual se realiza a través de la campimetría.

MANIFESTACIONES Y SÍNTOMAS La **BAJA VISIÓN** puede presentarse mediante distintas manifestaciones o síntomas, que se pueden agrupar en cuatro tipos:

- A. Pérdida de visión macular: Visión lateral.
- B. Pérdida de visión lateral: Visión en túnel.
- C. Pérdida de visión no localizada.
- D. Pérdida de definición o de agudeza visual.

MANIFESTACIONES	PATOLOGÍAS
Visión Lateral	Degeneración Macular
Visión en Túnel	Glaucoma
	Retinitis Pigmentaria
Pérdida de visión no localizada	Desprendimiento de Retina
	Retinopatía Diabética
Pérdida de agudeza visual	Cataratas
	Miopía Magna

Tabla 1.1
Manifestaciones de
Baja Visión junto
con sus patologías
asociadas.

PATOLOGÍAS En general, los síntomas de BAJA VISIÓN pueden aparecer por muchas causas: algunas enfermedades, malformaciones genéticas, la edad o traumatismos. Existen diferentes patologías que pueden causar BAJA VISIÓN, entre las que se encuentran la degeneración macular, el desprendimiento de retina, la retinosis pigmentaria, la retinopatía diabética, el glaucoma, las cataratas y la miopía magna. En la Tabla 1.1 se muestran las diferentes manifestaciones de BAJA VISIÓN junto con sus patologías asociadas.

*La RETINOSIS
PIGMENTARIA
también es
conocida como
RETINITIS
PIGMENTARIA O
PIGMENTOSA*

INCIDENCIA Según la OMS [75], la prevalencia de la BAJA VISIÓN está entorno al 2 – 3 % en los países desarrollados. La presencia en Europa de pacientes con BAJA VISIÓN está por encima de los 20 millones, de los cuales más de 2 millones son españoles [6]. En cuanto a los EEUU, la incidencia de glaucoma y de retinitis pigmentaria, dos de las enfermedades que pueden provocar VISIÓN EN TÚNEL, es de 2 pacientes por cada 100 adultos mayores de 40 años para el glaucoma, y de entre 20 y 33 individuos de cada 100 000 para la retinitis pigmentaria [64].

*VISIÓN EN TÚNEL:
Pérdida de la visión
lateral con visión
únicamente a través
de la fovea.*

Test de Snellen

Con el test de Snellen se determina el menor tamaño de detalle que es capaz de ver un paciente. En este test se solicita al paciente que indique cuál es la letra de menor tamaño que es capaz de identificar completamente de una tabla situada a 6 metros de distancia. El tamaño de las letras se construye partiendo de un tamaño base, que ha sido obtenido previamente midiendo el tamaño mínimo de las letras que son capaces de ver los pacientes con buena visión a 6 metros de distancia. Finalmente, la medida de la *agudeza visual* se indica con una fracción, donde el numerador es 6 y el denominador indica la distancia a la que un paciente con buena visión sería capaz de ver esa misma letra. Así pues, la *agudeza visual* de 6/6 (o 1,0) es el valor base en la que un paciente es capaz de ver perfectamente las letras de tamaño base a 6 metros de distancia. Por otra parte, si a 6 metros de distancia un paciente sólo puede ver las letras de mayor tamaño, la medida de *agudeza visual* es de 6/60 (o 0,1); esto es equivalente a decir que un paciente con buena visión sería capaz de ver ese tamaño de letra a 60 m. de distancia.

ACTUACIÓN Y REHABILITACIÓN Los pacientes con BAJA VISIÓN no pueden recuperar visión, es decir, que cuando ya se han agotado todas las posibilidades quirúrgicas, oftalmológicas y ópticas posibles, la única opción posible es una rehabilitación que les permita sobrellevar su enfermedad y mejorar su calidad de vida. Para esta rehabilitación se utilizan técnicas visuales que les permiten aprovechar al máximo el remanente visual que les queda. También se hace uso de elementos ópticos de aumento y de aparatos electrónicos de ayuda a la visión, sobre todo para la lectura.

APLICACIÓN CIENTÍFICO-TÉCNICA En general todos los pacientes con BAJA VISIÓN tienen graves dificultades para su propio manejo autónomo, por lo que la Comisión Europea, a través del 7º Programa Marco [35], ha impulsado el desarrollo de sistemas de apoyo para este colectivo. En particular, los pacientes que sufren de *visión en túnel* presentan algunas características que permiten el desarrollo de aplicaciones para la mejora de su propia autonomía. Al perder la visión lateral, carecen de todas las referencias espaciales del entorno, por lo que sólo son capaces de identificar aquello que está justo enfrente de ellos, ya que es precisamente esta zona la que está cubierta por la fóvea.

Al ser dicha zona la de mayor sensibilidad y agudeza visual del ojo, existen estudios [64, 113] en los que se proyecta en ella una imagen reducida de un campo visual mayor, de tal forma que el individuo con BAJA VISIÓN puede ver una miniatura de la escena que le rodea. Para que el paciente pueda observar también la imagen real que se le presenta de manera natural en su fóvea, la imagen del entorno en miniatura se presenta mediante un esbozo de bordes. Hay estudios [3, 65, 74] que han demostrado que el sistema visual humano es capaz de ignorar y acomodar la vista casi a la perfección ante la presencia de elementos no completamente opacos situados a diferentes niveles de profundidad. A este efecto se le denomina *ceguera no atencional*. Gracias a la cual, el paciente es capaz de ignorar la miniatura si lo que desea es observar la imagen natural captada por su fóvea, desenfocando dicha imagen en miniatura, y viceversa. Se puede observar dicho comportamiento ante manchas en un cristal: es posible enfocar la mancha presente en el cristal e ignorar la imagen real del fondo, o bien enfocar la imagen que hay en el fondo tras el cristal, ignorando la mancha.

En este campo de aplicación, un procesamiento eficaz de la escena, en el que se puedan eliminar las sombras y los efectos causados por una iluminación no uniforme de la imagen, permitiría mostrar imágenes de bordes en miniatura con mayor utilidad para los pacientes que las actuales, en las que las zonas en sombra no suelen estar bien representadas. También permitiría la utilización de técnicas más complejas de inteligencia artificial, que adapten las miniaturas a las necesidades del paciente en cada momento. Por ejemplo, el sistema debería ajustar su comportamiento según el usuario esté andando por la calle o leyendo: en el primer caso se requiere menos precisión y más rapidez, mientras que en el segundo, el requisito fundamental es la precisión, siendo la rapidez de respuesta menos importante.

1.2.2 *Sistemas inteligentes de Realidad Aumentada: Beyond Human Capabilities*

El término *Realidad Virtual* fue acuñado por Lanier [62] a principios de la década de 1980 para definir la generación de gráficos utilizados para crear escenas virtuales. En contraposición, la REALIDAD AUMENTADA mezcla imágenes reales con gráficos y/o información textual, ambos generados por una computadora. Hay que indicar que la información generada se obtiene a partir de las imágenes reales captadas por el sistema, es decir, que los gráficos añadidos no son independientes de las imágenes captadas, sino que se produce una inmersión de estos en las imágenes reales.

La Realidad Aumentada mezcla imágenes reales con otras generadas artificialmente.

ÁMBITO DE DEFENSA Los primeros sistemas de REALIDAD AUMENTADA vinieron de la inversión en investigación y desarrollo de los ejércitos y ministerios de defensa de diferentes países. Uno de los primeros fue el sistema de ayuda de navegación aeronáutica de los pilotos de combate [90, 91], en el que los pilotos llevan un casco con un visor transparente en el que se proyectan datos de pilotaje (compás, altímetro, velocímetro, etc.), de tal manera que no tienen que retirar la vista del frente para obtener dicha información. Por otro lado, actualmente están tomando mucha importancia los sistemas de exploración, soporte táctico y de ayuda balística para infantería. Ministerios de Defensa de distintos países están apoyando la investigación y el desarrollo de estos sistemas, junto con armas más seguras y equipamiento que proporciona un mayor grado de defensa al combatiente, analizados y desarrollados en el proyecto “*combatiente del futuro*” [36].

La REALIDAD AUMENTADA en defensa se utiliza para proporcionar información extra y mayor seguridad a los soldados en combate. Por ejemplo, dotando a los combatientes de sistemas que permitan el disparo desde posiciones ocultas, es decir, sin exponer al soldado a la visión directa del blanco. Otra funcionalidad deseada es la de proporcionar información táctica (sobreimpresión de imágenes de emisión de calor, distancia al blanco, posición de otros miembros del equipo de ataque, posibilidad de visualización de planos de servicio, topográficos, etc.). Para ello, las imágenes reales son captadas por cámaras de alta sensibilidad y otros sensores (que habitualmente están montados sobre el fusil de asalto o sobre el casco del soldado) y, tras realizar un procesamiento computacional de dichas imágenes, los resultados son presentados al soldado mediante gafas con microproyectores o de cristal líquido.

ÁMBITO MÉDICO En las intervenciones quirúrgicas hay una tendencia a realizar el mínimo número de incisiones externas y del menor tamaño posible. Para realizar la cirugía interna necesaria, se introducen por algunas de las incisiones unas pequeñas pinzas que permiten manipular los tejidos, vasos sanguíneos, etc., que estén involucrados en la operación. Como las operaciones quirúrgicas involucran la consulta de muchos datos sobre el estado del paciente, además del conocimiento médico para ser capaz de identificar los puntos de actuación, junto con los órganos, vasos sanguíneos, músculos y demás elementos implicados en la operación, la consulta de los datos puede implicar retirarse

de la zona de operación, lo que puede provocar que se deje de observar algún punto interior del paciente. Para solucionar este problema, actualmente también se realizan otras incisiones por las que se introducen unas microcámaras que permiten al cirujano observar el interior del paciente. Los cirujanos y el resto de personal sanitario observan las imágenes captadas por las microcámaras a través de monitores de visualización presentes en el quirófano. De esta forma, al ver la operación a través de un monitor, hay facilidad de incluir en la imagen mostrada por pantalla los datos relevantes sobre el paciente. En estas operaciones, también existe la posibilidad de marcar digitalmente elementos del paciente y que un ordenador realice el seguimiento de los mismos, para informar de anomalías a lo largo de la operación, pudiendo el cirujano centrarse prioritariamente en otros aspectos de la cirugía.

También dentro del ámbito médico, el procesamiento de imágenes médicas permite la reconstrucción 3D de tomografías, ecografías y demás sistemas de exploración no invasiva. La inclusión de esta información, previamente obtenida, en el momento de las operaciones permite a los cirujanos tener ayudas visuales virtuales del campo quirúrgico. En algunos casos, incluso se permite hacer rotaciones y cambiar el punto de vista, simulando cómo se vería el elemento que se está operando desde otra posición.

ÁMBITO CULTURAL Se pueden visitar los museos de diferentes maneras: simplemente observando los objetos expuestos, leyendo sus etiquetas explicativas, con un libro-guía que nos cuente todos los detalles, con un guía que además de explicarnos todos los datos de interés también atienda todas nuestras dudas, etc. Sin embargo, seguramente es preferible que los autores de los objetos o los protagonistas de las escenas que se están observando sean los que expliquen todas esas interioridades. En la mayoría de los casos eso no es posible, bien porque el autor o los protagonistas han fallecido, o porque simplemente el autor no puede estar siempre en dicho lugar para atender a los visitantes. En [11] se muestra cómo, mediante un sistema de REALIDAD AUMENTADA, se integran actores-guía en el cementerio histórico de Oakland, en Atlanta, que simulan ser “fantasmas” de la guerra civil americana. Estos “fantasmas” virtuales comentan el escenario real donde vivieron, explicando los acontecimientos que sucedieron durante su vida, etc. Para que “aparezcan estos fantasmas” durante la visita libre a dicho cementerio, los visitantes llevan un sistema de gafas que contiene un microproyector y unos auriculares incorporados, además de un GPS diferencial, un sistema de control inercial y sistemas de captura óptica. Mediante estos sistemas se determina la posición del visitante y se le proyecta una grabación personalizada y adaptada a lo que esté viendo (teniendo en cuenta características del terreno, etc.), de manera que sea plausible y coherente no sólo lo que explique el “fantasma”, sino su posicionamiento y su movimiento.

El enfoque anterior, aunque da muy buenos resultados en ambientes controlados, no es aplicable en todos los casos. Por ejemplo, las visitas guiadas de recintos o incluso de ciudades o países tienden a ser muy restrictivas y poco flexibles, presentando una serie de puntos en los que se describen algunos conceptos clave y dejando al usuario la misión de localizar los sitios de interés descritos. Los teléfonos actuales,

permiten solucionar este problema, gracias a que incluyen una multitud de sensores (cámara, GPS, giróscopos, acelerómetros, brújulas, etc.) y procesadores de relativamente alta capacidad de cómputo. Ejemplos de este tipo son las aplicaciones *Layar Reality Browser* [63] o *Wikitude* [105]. Éstas, utilizando la cámara del teléfono como entrada de imágenes y la señal del GPS integrado junto con la brújula para seguir la localización del usuario, permiten procesar dichas imágenes, y señalar e identificar en ellas mediante etiquetas los diferentes monumentos, edificios singulares u otros aspectos de interés para el usuario, que se muestran en la pantalla del teléfono según se les va enfocando con la cámara integrada.

ÁMBITO DEPORTIVO Y TELEVISIVO En las retransmisiones deportivas, los espectadores desean tener muchos datos sobre el juego que están viendo. Así, las televisiones proporcionan información visual sobreimpresionada sobre las imágenes reales en las que, dependiendo del deporte, se muestran distancias, estadísticas de jugadores o líneas dinámicas del juego (como pueden ser líneas de fuera de juego, líneas de pase, líneas de “primer down”, zonas libres de marcaje, líneas de ataque, estructuras defensivas, etc.). También se suele incluir publicidad sobreimpresionada como si estuviese pintada sobre el propio campo de juego. Todos estos datos se muestran sin que la presencia de los mismos oculten la jugada. Para ello, las televisiones calibran sus cámaras de manera que antes de la retransmisión deportiva ya conocen las proporciones del terreno de juego, y consecuentemente tienen una representación tridimensional en la que pueden adaptar todos los gráficos al perfil real. Además, para que los cálculos parezcan inmediatos, se retrasa la señal de vídeo en unas decenas de fotogramas, lo que permite alterar dichos fotogramas añadiendo elementos gráficos generados por ordenador.

Por otro lado, para los entrenadores y analistas deportivos el uso del vídeo tiene mucha importancia, ya que con él pueden revisar las actuaciones propias, detectando errores, acciones de mayor efectividad, etc. También les permite estudiar a los equipos contrarios, tanto en ataque como en defensa, repasar acciones especiales, movimientos tácticos de jugadores en particular, etc. En este contexto, los sistemas informáticos permiten la automatización de algunas tareas, como por ejemplo, el seguimiento de un determinado jugador o del balón, y gracias a estos seguimientos automatizados, el ordenador puede trazar la ruta que han seguido los jugadores y facilitar su estudio.

ÁMBITO LÚDICO-RECREATIVO Las estimaciones indican que en el año 2011, el negocio de los videojuegos moverá al año unos 48.900 millones de dólares (más de 37.300 millones de euros) en todo el mundo [92]. Este volumen económico tan elevado, hace que esta industria sea una de las más pujantes dentro del ámbito de la informática, con grandes avances año tras año para mantenerse en la cresta de la ola de las demandas de los usuarios. En los videojuegos actuales, existe una demanda de gráficos hiperrealísticos y con controles de movimiento lo más naturales posibles. Esto se ha conseguido gracias al uso de tecnologías de REALIDAD AUMENTADA en las videoconsolas utilizando cámaras y mandos con diferentes sensores, tales como acelerómetros, girós-

copos, etc. Este proceso se generalizó en el 2003 con la introducción del sistema Eye Toy en la Sony PlayStation 2, aunque ya antes otros sistemas habían propuesto mecanismos parecidos para la generación de sistemas de REALIDAD AUMENTADA. En este tipo de videojuegos, se captan imágenes reales del usuario y su entorno mediante cámaras incluidas en las diferentes plataformas de juego, que son incluidas dinámicamente en el videojuego para proporcionar más realismo, de forma que el propio usuario es el que determina la respuesta del juego con los movimientos de su cuerpo. La generalización de las videoconsolas portátiles con cámaras, PSP y Nintendo DSi, ha permitido la integración de escenarios reales en los juegos, consiguiendo imbuir a los usuarios en juegos que transcurren en su entorno real. El videojuego *Invizimals*, publicado durante el año 2009, es un claro ejemplo de las posibilidades lúdicas de la REALIDAD AUMENTADA. En este videojuego desarrollado para la consola portátil PSP, el jugador debe atrapar unos “monstruos” virtuales, sólo visibles utilizando la cámara de dicha videoconsola. Además, se utilizan unas trampas (unos dispositivos físicos con forma cuadrada) que se colocan físicamente en las diferentes habitaciones y el jugador debe conseguir que los *Invizimals* caigan en dichas trampas.

1.3 MARCO DE TRABAJO

Tras plantear la motivación y unos posibles campos de aplicación de esta Tesis, a continuación se acota el marco de trabajo en el que se ha desarrollado la investigación. Las restricciones propuestas con respecto al tiempo de procesamiento han determinado la solución escogida. Por otra parte, el grado de calidad de la respuesta esperada también ha influido decisivamente en los métodos y en los parámetros de funcionamiento seleccionados.

1.3.1 *Esquemas de evaluación de los resultados del procesamiento de imágenes*

El principal aspecto que ha condicionado el desarrollo de esta Tesis Doctoral ha sido el de la búsqueda del método que proporcione una mejor calidad en su respuesta visual. Como se ha enunciado brevemente con anterioridad, y como se describirá más adelante, la evaluación de la calidad del resultado de las imágenes procesadas, en la mayoría de los casos, es una cuestión subjetiva. Esta característica ha marcado en gran medida las diversas decisiones que se han ido tomando, tanto en la investigación como en el desarrollo de las técnicas y de las herramientas propuestas en esta memoria.

1.3.2 *Procesamiento de imágenes en tiempo real*

Una aplicación de procesamiento de imágenes en tiempo real debería ser capaz de capturar las escenas y procesarlas con una tasa de entre 25 y 30 fotogramas por segundo y con un retardo menor de 1 fotograma; es decir, se debería procesar un nuevo fotograma cada 33 o 40 milise-

gundos, según la tasa de captura. Con esta frecuencia de procesamiento, el resultado se vería continuo, suave, y se presentaría coherente con la visión real. Sin embargo, el trabajo que se presenta en esta Tesis es una primera aproximación a la solución, por lo que no se ha impuesto como requisito imprescindible la velocidad de procesamiento, aunque sí se ha tenido en cuenta para el desarrollo de los algoritmos asociados a los métodos propuestos, seleccionando aquellas implementaciones de menor complejidad computacional.

1.4 PLANTEAMIENTO INICIAL

Teniendo en cuenta todos los requisitos, factores implicados y restricciones expuestos en las secciones anteriores, a continuación se expone el planteamiento inicial del trabajo de la presente Tesis Doctoral. Esta sección, que sirve como resumen de lo expuesto, y muestra las primeras decisiones de diseño de entre las opciones anteriormente descritas, se ha usado como punto de inicio de la investigación.

1.4.1 Problemática

La problemática que se desea resolver es la siguiente:

En general, los sistemas de visión artificial basados en la extracción de bordes son muy sensibles ante los cambios locales de iluminación en las escenas: Presentan baja capacidad de detección de bordes en las zonas en sombra de las escenas.

1.4.2 Ámbito de actuación

De todos los posibles ámbitos de actuación descritos con anterioridad, este trabajo se centra en:

El soporte visual para los pacientes con BAJA VISIÓN, en especial, para aquellos con sintomatología de VISIÓN EN TÚNEL.

1.4.3 Objetivos científicos de la Tesis

La presente Tesis Doctoral presenta varios objetivos que se pueden dividir entre objetivos primarios y objetivos secundarios. Los objetivos primarios son los que han marcado las directrices generales de la investigación y pretenden dar respuesta a las preguntas: “¿Qué debe realizar el sistema?” y “¿Cómo de buenos son los resultados que proporciona?” Por otra parte, los objetivos secundarios son aquellos que se desea que cumpla el sistema de manera adicional a los primarios. Con estos ob-

jetivos se pretende matizar la respuesta de “¿Cómo deben realizarse las acciones del sistema?”

OBJETIVOS PRIMARIOS

Como OBJETIVOS PRIMARIOS se desea:

- A. Desarrollar un método eficiente y robusto para obtener un esbozado de escenas reales con sombras.
- B. Obtener una evaluación de la calidad de las imágenes procesadas, obtenidas con diferentes métodos.

OBJETIVOS SECUNDARIOS

También se desea alcanzar los siguientes OBJETIVOS SECUNDARIOS:

- A. Obtener imágenes procesadas con buena calidad subjetiva para pacientes con BAJA VISIÓN.
- B. Optimizar el algoritmo de procesamiento para alcanzar una velocidad adecuada a la aplicación propuesta.
- C. Generar un mecanismo para la evaluación subjetiva de imágenes procesadas mediante encuestas de opinión.

1.4.4 Propuesta inicial

A partir de los OBJETIVOS PRINCIPALES se ha determinado qué herramientas y métodos se deben desarrollar. Para conseguir dichos objetivos ha sido necesario investigar y desarrollar un método de extracción de características que tenga un comportamiento invariante ante cambios de iluminación, así como la construcción de un sistema Web para la evaluación de la calidad de las imágenes procesadas mediante encuestas, aspecto que se ha considerado de interés para el cumplimiento del OBJETIVO SECUNDARIO C.

Método para la extracción invariante de características ante cambios de iluminación

Se conoce que el paradigma LIP permite extraer bordes de manera invariante ante cambios de iluminación mediante una modificación del método de Sobel [29], aunque esta aproximación presenta problemas de generalización a otros métodos de extracción de bordes más complejos. Sin embargo, la principal ventaja que presenta es que la extracción de bordes se realiza de manera simultánea al mecanismo que lo hace invariante ante los cambios de iluminación, frente a otras estrategias que realizan esto en varias etapas, aumentando el tiempo de computación y reduciendo, por tanto, la velocidad de proceso. Así pues, se ha decidido investigar este paradigma para conseguir un mecanismo generalizable que permita aplicar otros mecanismos de extracción de bordes con la mínima dificultad de adaptación de los mismos.

LIP: Del inglés, LOGARITHMIC IMAGE PROCESSING. Ver Cap. 2.2

Sistema Web para la evaluación subjetiva de imágenes mediante encuestas

Dada la complejidad que supone reunir a suficientes pacientes con BAJA VISIÓN en un mismo lugar y un mismo día para que puedan contestar las encuestas de opinión, se ha pensado que Internet puede proporcionar un medio ideal para resolver esta problemática, ya que permite que los pacientes puedan acceder al cuestionario remotamente sin que se tengan que desplazarse a un lugar concreto para rellenarlo. Añadiendo sistemas de almacenamiento (como bases de datos) en el servidor se permitiría que los pacientes accediesen en cualquier momento, e incluso se permitiría a un paciente continuar una encuesta no terminada completamente. Esto último puede ser muy habitual, debido al cansancio visual que el esfuerzo de rellenar una encuesta puede provocar en este tipo de pacientes.

1.5 ESTRUCTURA DE LA TESIS

Esta Tesis Doctoral se ha estructurado en cinco partes, que a su vez se han dividido en capítulos, y estos en secciones y subsecciones. La composición de las diferentes partes es la siguiente:

INTRODUCCIÓN En esta Parte I se expone la problemática que esta Tesis pretende resolver. También se plantean, desde un punto de vista científico, las posibilidades que la ciencia y la tecnología actual ofrecen al desarrollo de sistemas que resuelvan la problemática expuesta. Se incluyen ejemplos de distintos ámbitos de actuación en los que esta problemática está presente. Para finalizar el primer capítulo, se fijan los OBJETIVOS y una primera aproximación a la solución. Se continúa con una revisión de técnicas de procesamiento de imágenes, exponiendo diversos métodos de extracción de contornos. También se expone el paradigma LIP con todos sus operadores, rangos de funcionamiento, leyes, etc. Tras lo cual, se describen los mecanismos de evaluación de contornos existentes en el ámbito científico, tanto de tipo objetivo como subjetivo. En el siguiente capítulo se tratan en profundidad las

encuestas, describiendo los diferentes tipos que existen, las limitaciones que éstas presentan, el diseño genérico de los cuestionarios, los errores y sesgos más habituales y, finalmente, los métodos de validación de las mismas. Esta primera parte finaliza con una revisión de los cuestionarios para evaluación de los pacientes con BAJA VISIÓN, en la que se analizan los dominios comunes más relevantes.

MÉTODOS Y MATERIALES La Parte II de la Tesis Doctoral consta de tres capítulos. En el primero de ellos, se presenta el nuevo operador *Convolución-LIP* con dos implementaciones diferentes, una de las cuales es una optimización que tiene en cuenta la eficiencia computacional de las operaciones involucradas. En el siguiente capítulo de esta Parte, se expone la HIPÓTESIS DE TRABAJO y se fijan los objetivos funcionales que se desean evaluar a través de encuestas de opinión y los métodos de extracción de contornos que se van a evaluar. Esto llevará a determinar los DOMINIOS de trabajo de las mismas y las *tareas de visión* asociadas a las preguntas de los cuestionarios. Se finaliza dicho capítulo con una selección de imágenes de cada uno de los DOMINIOS, con los que se obtendrán los diferentes mapas de contornos para cada algoritmo seleccionado, que serán evaluados a través de encuestas. El tercer y último capítulo de esta Parte de la Tesis incluye un compendio de recomendaciones para el desarrollo de cuestionarios para la evaluación subjetiva de la calidad del procesamiento de contenidos visuales. Este capítulo concluye con una implementación de un sistema Web que proporciona el soporte para el diseño y la respuesta de encuestas de evaluación de calidad de contenidos multimedia.

RESULTADOS EXPERIMENTALES Esta tercera parte confirma el cumplimiento de la HIPÓTESIS DE TRABAJO a través de experimentos y encuestas de investigación. En el primero de los dos capítulos que componen dicha Parte III, se muestra el funcionamiento del nuevo operador *Convolución-LIP* con diferentes algoritmos de *extracción de contornos* y de *emborronamiento de imagen*. Se comparan los diferentes algoritmos en las diferentes implementaciones dentro del paradigma LIP y, en algunos casos, también se comparan con los métodos No LIP de los algoritmos evaluados, para mostrar la supremacía de la propuesta frente a las otras versiones. Estas comparaciones se realizan a nivel visual, aunque también se incluyen experimentos que evalúan el tiempo de cómputo y la ganancia de rendimiento relativo entre ellos. En el siguiente capítulo, se muestra el proceso completo de la evaluación mediante cuestionarios de opinión de los algoritmos que se seleccionaron en la Parte II. En dicho capítulo, se muestran los resultados obtenidos a partir de las respuestas de los usuarios con un grado de VISIÓN ESTÁNDAR y los que presentan problemas de BAJA VISIÓN para cada algoritmo evaluado. Además de estos resultados finales, se muestra cómo se han obtenido los cuestionarios, la población que ha respondido la encuesta, etc.

DISCUSIÓN Y CONCLUSIONES Esta parte está compuesta por dos capítulos, en los que se exponen las discusiones de los resultados finales, las conclusiones y las principales aportaciones científicas que propor-

ciona esta Tesis Doctoral. También, se indican posibles vías de investigación futura.

BIBLIOGRAFÍA Y ANEXOS Se incluyen todas las referencias utilizadas a lo largo de la Tesis. Finalmente, se añaden una serie de anexos sobre demostraciones matemáticas de algunos operadores del paradigma LIP y se incluyen las preguntas e imágenes que se han usado en los cuestionarios de evaluación utilizados en esta investigación.

ÍNDICE

2.1	Modelos de formación de imágenes	22
2.1.1	Filtrado homomórfico	22
2.2	Descripción del modelo LIP	24
2.2.1	Nomenclaturas y rangos	25
2.2.2	Bases matemáticas de LIP	26
2.2.3	Relación con otros modelos logarítmicos	29
2.2.4	Leyes físicas y psico-físicas que cumple LIP	30
2.3	Modelos de ruido en imágenes	33
2.3.1	Ruido en imágenes	33
2.3.2	Emborronamiento de imágenes	34
2.4	Extracción de bordes	35
2.4.1	Filtros basados en el gradiente de la imagen	36
2.4.2	Filtros basados en el Laplaciano de la imagen	38
2.4.3	Método de Canny para la extracción de bordes	41
2.5	Conclusiones	43

A lo largo de esta Tesis Doctoral se utilizan conceptos y métodos de PROCESAMIENTO DE IMÁGENES que deben ser explicados más a fondo para una mejor comprensión de los algoritmos y métodos desarrollados. En este capítulo se hace una revisión de las principales técnicas de extracción de bordes en imágenes. También se muestra un estudio actualizado de los métodos que otros investigadores han utilizado para eliminar o minimizar la dependencia de los cambios de iluminación en diversos algoritmos de procesamiento de imágenes. La gran mayoría de estos métodos se desarrollaron principalmente para abordar el problema de las sombras en las imágenes, cuya presencia (en distintos escenarios) suele afectar negativamente a los algoritmos de procesamiento.

Este capítulo comienza con la descripción del modelo de formación de imágenes, tras lo cual se describen algunos modelos de ruido de generación natural. Tras esto, se describe el modelo LIP, que se servirá como base matemática para la construcción de los algoritmos que se propondrán en esta Tesis. A continuación, se muestran diversos métodos para la extracción de los contornos de los objetos, basados en operadores que utilizan el gradiente y/o el laplaciano de las imágenes, obtenidos utilizando el operador matemático *convolución*.

2.1 MODELOS DE FORMACIÓN DE IMÁGENES

El modelo multiplicativo de formación de imágenes [42, 78] afirma que las imágenes captadas son una superposición de longitudes de onda de diferentes frecuencias que responden, en una primera aproximación, a las componentes de la iluminación de la imagen y a la luz reflejada [33] (o transmitida, dependiendo del tipo de imagen considerada) por los objetos presentes en la escena que se está observando. A estas dos componentes se les denomina respectivamente, *iluminación* y *reflectancia* o *transmitancia*, dependiendo del tipo de imagen que se esté considerando (ver cuadro ¿REFLECTANCIA O TRANSMITANCIA? al pie de página). Este modelo sigue la siguiente formulación:

$$f(x, y) = i(x, y) \cdot r(x, y), \quad (2.1)$$

donde $f(x, y)$ es el valor de la función de intensidad recibida por el observador asociado al punto (x, y) del sistema de referencia del propio observador. Este valor se define mediante la multiplicación de dos factores: la *iluminación* ($i(x, y)$) de la imagen que se recibe en dicho punto y la componente de *reflectancia*, $r(x, y)$, que representa el valor de la proyección del objeto asociado a la posición (x, y) .

2.1.1 Filtrado homomórfico

Habitualmente, se desea poder procesar imágenes evitando la influencia de la *iluminación*, ya que ésta puede no ser uniforme, es decir, puede haber sombras, brillos, etc., que podrían desvirtuar el contenido real (los objetos presentes) de la escena observada. Esto significa que se desea descartar toda la información de la componente de *iluminación* y trabajar únicamente con la componente de *reflectancia*; sin embargo, las dos componentes se entremezclan entre sí espacialmente de tal manera que no es posible separarlas directamente mediante operadores lineales.

¿Reflectancia
o Transmitancia?

Las imágenes generadas mediante transmisión de luz son llamadas imágenes de TRANSMITANCIA, entre las que se encuentran por ejemplo las de microscopía. En éstas existe un foco luminoso cuyo haz de luz traspasa el objeto observado, que no es totalmente opaco, y el receptor recibe información visual de la cantidad de luz que los objetos observados permiten traspasar.

En contraposición, existen las llamadas imágenes de REFLECTANCIA, que son las imágenes captadas por los observadores debidas a la reflexión de la luz de un foco sobre las superficies de objetos opacos y no especulares. Lo que el observador recibe en este caso es la cantidad de luz que los objetos reflejan. En general, todas las fotografías realizadas con cámaras de exposición son de este tipo.

Las imágenes que observamos son una mezcla multiplicativa de iluminación y reflectancia.

Utilizando el espacio transformado de Fourier se puede mostrar que la *iluminación* afecta a todas las frecuencias, ya que la transformada de Fourier de la multiplicación de las dos componentes no es separable [42], es decir:

$$\mathcal{F}\{f(x, y)\} \neq \mathcal{F}\{i(x, y)\} \cdot \mathcal{F}\{r(x, y)\}, \quad (2.2)$$

donde la transformada de Fourier viene indicada por $\mathcal{F}\{\cdot\}$.

La no separabilidad de estas dos componentes impide la eliminación directa de la *iluminación*. Sin embargo, se ha podido demostrar [77, 78] que:

$$z(x, y) = \ln(f(x, y)) = \ln(i(x, y)) + \ln(r(x, y)) \quad (2.3)$$

Por lo tanto,

$$\mathcal{F}\{z(x, y)\} = \mathcal{F}\{\ln(f(x, y))\} = \mathcal{F}\{\ln(i(x, y))\} + \mathcal{F}\{\ln(r(x, y))\} \quad (2.4)$$

La ecuación 2.4 muestra la base del modelo **MHIP**. La teoría de probabilidades [41] afirma que si tanto $\ln(i(x, y))$ como $\ln(r(x, y))$ son dos funciones estadísticamente independientes, el histograma del logaritmo de la combinación de ambas se obtiene mediante la convolución de los histogramas del logaritmo de cada una de las componentes. En la mayoría de los casos, no se suele poder acceder a la fuente de *iluminación*, por lo que hay que hacer una estimación de dicha componente. A priori, para un conjunto de objetos con superficies Lambertianas [56] se puede suponer que la *iluminación* varía suavemente, y de hecho, la *iluminación* es una componente que se localiza principalmente en las bajas frecuencias. Por tanto, se puede afirmar que para un entorno de tamaño reducido es prácticamente constante. Por ello, se puede aceptar como aproximación de primer orden que:

$$\text{LPF}\left[\mathcal{F}\{\ln(f(x, y))\}\right] \approx \mathcal{F}\{\ln(i(x, y))\}, \quad (2.5)$$

donde $\text{LPF}[\cdot]$ (siglas en inglés de **LOW PASS FILTER**, *Filtro Paso Baja*) es una función que calcula el filtrado paso baja en frecuencia de una función dada. Con lo que para pequeños entornos de un punto, se puede asumir que:

$$f(x, y) = i \cdot r(x, y) \quad (2.6)$$

Por lo que, sustituyendo (2.5) en (2.4), se tiene que:

$$\mathcal{F}\{\ln(f(x, y))\} = \text{LPF}\left[\mathcal{F}\{\ln(f(x, y))\}\right] + \mathcal{F}\{\ln(r(x, y))\} \quad (2.7)$$

De (2.7) se puede calcular $\ln(f(x, y))$ puesto que se tiene la imagen completa. Además se puede estimar $\ln(i)$ mediante algún filtro paso baja (ya que como se ha indicado anteriormente, $\ln(i)$ se encuentra principalmente en las bajas frecuencias), por lo que finalmente se puede calcular la componente de *reflectancia*:

$$\mathcal{F}\{\ln(r(x, y))\} = \mathcal{F}\{\ln(f(x, y))\} - \text{LPF}\left[\mathcal{F}\{\ln(f(x, y))\}\right] \quad (2.8)$$

MHIP: Del inglés, MULTIPLICATIVE HOMOMORPHIC IMAGE PROCESSING, procesamiento de imágenes multiplicativas homomórficas. Las superficies Lambertianas reflejan la luz en todas direcciones, de manera independiente al ángulo de incidencia.

También se puede calcular mediante:

$$\mathcal{F}\{\ln(r(x, y))\} = \ln\left(\frac{\mathcal{F}\{f(x, y)\}}{\text{LPF}\left[\mathcal{F}\{f(x, y)\}\right]}\right) \quad (2.9)$$

Para reducir el coste computacional, habitualmente se utiliza (2.8), ya que es más sencilla de calcular que (2.9). Por este mismo motivo, de manera general no se realiza el cálculo de la transformada de Fourier, por lo que la fórmula final queda:

$$\ln(r(x, y)) = \ln(f(x, y)) - \ln(i) \quad (2.10)$$

donde $\ln(i)$ representa la *iluminación* de un determinado entorno de un punto, obtenido como el filtro paso baja del logaritmo de la imagen.

En la Sec. 2.2 se describe otro modelo de construcción de imágenes, que permite un manejo más eficiente (en el sentido de robustez) de las componentes de *iluminación* y de *reflectancia* (o *transmitancia*) que el modelo MHIP.

2.2 DESCRIPCIÓN DEL MODELO LIP

Fue a finales de la década de los 80 y a lo largo de la década de los 90, cuando Jourlin y Pinoli [58, 59] expusieron un nuevo paradigma para el procesamiento de imágenes por proyección de luz (como es el caso de imágenes de microscopía). El estudio y posterior desarrollo de LIP se realizó para obtener una respuesta satisfactoria a la pregunta “¿Cómo evita el sistema visual humano el problema del resultado fuera de rango al sumar dos imágenes?”. Pretendían dar solución a un problema tan habitual en los sistemas de procesamiento de imágenes digitales, como es el del desbordamiento del rango en la adición de dos imágenes o en el ajuste multiplicativo por una constante. Posteriormente, otros autores hicieron grandes contribuciones a este paradigma a lo largo de toda la década de los 90.

La aplicación de operadores tradicionales a imágenes digitales de tipo natural (obtenidas tanto por proyección o transmisión de la luz, como por su reflejo en los objetos de la escena) no es una solución matemáticamente robusta ni tiene una justificación física o psico-física [86], así que los autores originales del paradigma LIP describieron operadores matemáticos que, además de realizar la combinación y la amplificación (desde el punto de vista del procesamiento) de imágenes de microscopía, también eran consistentes con la naturaleza física de las imágenes que se trataban.

En este ámbito, los autores comprobaron que al superponer en el microscopio dos platinas con dos muestras, el observador puede observar la composición de ambas muestras. Es decir, que al sumar dos imágenes el observador obtiene una imagen dentro del rango visible. También, pudieron observar que aunque aumentasen la iluminación incidente en la platina, la imagen observada continuaba siendo visible. Por tanto, no es posible obtener ninguna combinación de muestras y de iluminación que proporcione un resultado que no esté dentro del

rango visible por el ojo humano (es decir, combinando imágenes en el rango visible no es posible obtener una imagen ultravioleta o infrarroja). Tomando como punto de partida estas observaciones, los autores definieron el brillo de un punto de una imagen en el modelo LIP, como la cantidad de luz que pasa por un *filtro lumínico* con una determinada *función de absorción* (o el grado de opacidad que presenta dicho filtro lumínico en un punto).

2.2.1 Nomenclaturas y rangos

Antes de explicar el planteamiento matemático de LIP, se especificarán los rangos de valores y la notación de las diferentes variables y funciones que se utilizarán en este trabajo.

Para empezar, cualquier escalar se notará mediante letras griegas minúsculas ($\alpha \in \mathbb{R}$). Los vectores de escalares reales se notarán con letras minúsculas en negrita, como **a**, **b**. Como excepción a la regla anterior, notaremos como M al valor máximo del rango permitido en función de la profundidad de bits de la paleta utilizada. Dicho valor máximo es inalcanzable. El mínimo valor (también inalcanzable) permitido en las imágenes es el 0. Estas dos afirmaciones anteriores se justificarán de manera más extensa en la Sec. 2.2.4. De manera breve, la elección de estos extremos encuentra su justificación en el mundo real, en el que la oscuridad absoluta (sin incidencia de ningún fotón) o la iluminación máxima absoluta (inyección de infinitos fotones) en condiciones naturales no puede ocurrir. En el caso de su aparición en imágenes digitales, se produce debido al proceso de captura, cuantización y digitalización de la imagen real. Para solucionar la presencia de estos valores límite no admisibles por LIP, se puede realizar un redondeo al valor válido más cercano sin pérdida de generalidad.

Aunque originalmente LIP fue diseñado para utilizarse con imágenes, al tener una base matemática consistente y robusta, permite su generalización y uso a otros elementos, sin que estos tengan que representar necesariamente una imagen. Los autores denominan a estos elementos como *funciones de tonalidades de gris*. Éstas se nombrarán con letras minúsculas, por ejemplo, f y g , con rangos $f, g \in (0, M) \subseteq \mathbb{R}$. De manera adicional, las imágenes habitualmente utilizadas con LIP suelen tener una paleta de tonalidades de gris, que en esta Tesis Doctoral se notarán con letras mayúsculas, I , J , etc. El rango de estas imágenes será $I \in (0, M) \subseteq \mathbb{N}$, o de manera equivalente, $I \in [1, M - 1] \subseteq \mathbb{N}$. En general, todo lo que se afirme para las *funciones de tonalidades de gris*, también es cierto para las *imágenes con paleta de tonalidades de gris*, de manera que se pueden utilizar ambos términos de manera indistinta. Los filtros bidimensionales también se notarán con letras mayúsculas, por ejemplo, F .

Para poder utilizar el paradigma LIP, es necesario utilizar imágenes adaptadas a dicho modelo, no pudiéndose utilizar de manera directa las imágenes con la paleta habitual. Cada valor de la paleta de estas imágenes se denomina *tono de gris de la función de tonalidades de gris* (a partir de ahora se denominarán simplemente *tonos de gris*, por brevedad) y se representarán con un acento circunflejo sobre la misma letra, por ejemplo, $\hat{f} = M - f$, con $\hat{f} \in (0, M) \subseteq \mathbb{R}$. Para poder aprovechar la

Si trabajamos con imágenes de 8 bits de profundidad, el máximo valor será 255, con lo que, $M = 256$.

potencialidad matemática que brindan las estructuras algebraicas de los espacios vectoriales, se ha expandido matemáticamente el rango de trabajo de los *tonos de gris* a $\hat{f} = (-\infty, M) \subseteq \mathbb{R}$. Hay que tener en cuenta que sólo los *tonos de gris* positivos tienen una asociación física con las imágenes reales, mientras que los *tonos de gris* pertenecientes al rango $(-\infty, 0)$ tienen sólo un significado matemático y no real.

Como se mostrará en la Sec. 2.2.2, LIP permite trabajar de dos maneras diferentes, unas de las cuales incluye una función isomórfica de transformación de los elementos. Así pues, los *tonos de gris* transformados utilizando la función isomórfica de LIP, por abreviar a partir de ahora, *tonos de gris* LIP, se notarán con sus respectivas letras con una virgulilla ($\tilde{}$) sobre la misma; por ejemplo, $\tilde{f} \in \mathbb{R}$.

2.2.2 Bases matemáticas de LIP

Los autores propusieron una estructura que, además de resolver la problemática de las operaciones fuera de rango, cumple muchas leyes físicas, psico-físicas y de percepción visual humana. Existen dos “mecanismos” para aplicar LIP a cualquier técnica de procesamiento de imágenes:

- Usar imágenes “originales” y un conjunto de operadores “modificados”.
- Usar imágenes “transformadas” y los operadores “tradicionales”.

A continuación se mostrará una breve descripción de cada uno de estos dos mecanismos indicados previamente.

Método Directo: Imágenes “Originales” y operadores “modificados”

Esta primera opción, que se ha dado en llamar *directa* por el uso directo de las imágenes originales, se genera mediante un espacio vectorial, que está definido por:

- Un conjunto de imágenes con valores de *tonos de gris*, notado como G^+ , que son imágenes “usuales” a las que se les ha aplicado una inversión de la escala:

$$\hat{f} = M - f \quad (2.11)$$

- Una operación suma modificada, \triangle , definida como:

$$\hat{f} \triangle \hat{g} = \hat{f} + \hat{g} - \frac{\hat{f} \cdot \hat{g}}{M} \quad (2.12)$$

Este operador garantiza que la suma de dos *tonos de gris* cualesquiera será, a su vez, otro *tono de gris* (y por tanto, se puede asegurar que el resultado se mantendrá dentro de rango). La demostración matemática de esta propiedad se muestra en el Apéndice A. Visualmente, se puede apreciar el resultado de esta operación en la Fig. 2.2, en la que se muestra la suma de las dos imágenes de la Fig. 2.1, tanto en su versión estándar como en su versión



Figura 2.1
Muestra de las dos
imágenes de ejemplo.



Figura 2.2
Suma de dos imágenes. Izquierda, suma estándar. Derecha, suma LIP. En rojo, todos los píxeles que se encuentran fuera de rango.

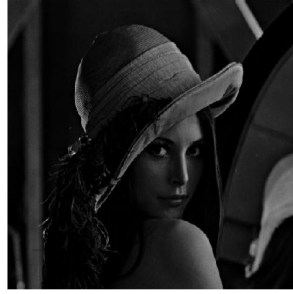


Figura 2.3
Multiplicación escalar de una imagen por 4. Izquierda, multiplicación estándar. Derecha, multiplicación LIP. En rojo, todos los píxeles que se encuentran fuera de rango.

LIP. Como se muestra en la imagen de la izquierda de la Fig. 2.2, existen muchos píxeles cuyo valor está fuera de rango (señalados en rojo), mientras que en la imagen de la derecha de la Fig. 2.2, todos los píxeles contienen valores dentro de rango.

- Un operador que implementa una multiplicación modificada de un *tono de gris* por un escalar, Δ , definida por:

$$\alpha \Delta \hat{f} = M - M \cdot \left(1 - \frac{\hat{f}}{M}\right)^\alpha \quad (2.13)$$

Este operador garantiza que la multiplicación de un escalar por un *tono de gris* es otro *tono de gris*, es decir, el resultado está definido dentro del mismo rango de funcionamiento. La demostración de esta propiedad que presenta este operador se muestra en el Apéndice B. En la Fig. 2.3 se muestra el resultado de la multiplicación de una de las imágenes de ejemplo de la Fig. 2.1 por 4. En la multiplicación “clásica” (imagen izquierda de la Fig. 2.3), muchos píxeles de la imagen resultado se han desbordado y poseen valores fuera de rango, mientras que en la imagen resultado de la

multiplicación LIP (imagen derecha de la Fig. 2.3) no hay ningún píxel con valor fuera de rango.

- Finalmente, para extender la estructura algebraica que proporcionan $(\mathbb{G}^+, \Delta, \Delta)$ a un espacio vectorial, se define de manera matemática el inverso de un *tono de gris*, es decir, se define la parte negativa de \mathbb{G}^+ , que se notará como \mathbb{G}^- . Se implementa un nuevo operador, $\hat{\Delta}$ que define la resta de dos *tonos de gris* como:

$$\hat{f} \hat{\Delta} \hat{g} = M \cdot \frac{\hat{f} - \hat{g}}{M - \hat{g}} \quad (2.14)$$

Partiendo de la idea expresada en la Sec. 2.2 en la que, en el modelo LIP, el concepto de brillo de un punto de una imagen se define como la cantidad de luz que permite pasar un *filtro lumínico* con un cierto grado de opacidad en dicho punto. Tomando el 0 como el máximo nivel de iluminación posible (debido a la inversión del rango de los *tonos de gris* indicado en el primer ítem), los *tonos de gris* $\hat{f} \in \mathbb{G}^-$ no tienen significado asociado en el mundo real. Puesto que una *función de tonos de gris* negativa, significaría que hay un punto en el filtro que no sólo no es opaco y por tanto, no reduce la cantidad de luz incidente que permite pasar, sino que incrementa la iluminación del punto, pudiendo incluso ser más brillante que el valor máximo posible.

Gracias a este nuevo conjunto, $\mathbb{G} = \mathbb{G}^+ \cup \mathbb{G}^-$, se puede definir un espacio vectorial compuesto por $(\mathbb{G}, \Delta, \Delta, \Delta)$.

Utilizando los operadores descritos, se han propuesto otros [29, 85, 86]. En este trabajo se utilizará el *Sumatorio-LIP*, que fue diseñado partiendo de la *Suma-LIP* especificada con anterioridad:

$$\hat{\Delta}_{i=1}^n \hat{f}_i = \hat{f}_1 \hat{\Delta} \hat{f}_2 \hat{\Delta} \dots \hat{\Delta} \hat{f}_n \quad (2.15)$$

Método mediante la “Transformada”: Imágenes “Transformadas” y operadores “tradicionales”

La segunda opción es transformar las imágenes, trabajar utilizando los operadores “tradicionales”, y finalmente, restaurar la imagen resultante al espacio original mediante la inversa de la transformada inicial.

La transformada se realiza mediante una función llamada *transformada isomórfica* que se define como:

$$\tilde{f} = \varphi(\hat{f}) = -M \cdot \ln \left(1 - \frac{\hat{f}}{M} \right) \quad (2.16)$$

La inversa de la función de transformación se denomina *transformada isomórfica inversa* y se define como:

$$\hat{f} = \varphi^{-1}(\tilde{f}) = M \cdot \left(1 - e^{-\frac{\tilde{f}}{M}} \right) \quad (2.17)$$

2.2.3 Relación con otros modelos logarítmicos

Existen otros paradigmas, definidos con anterioridad a LIP, que resuelven el problema de las operaciones fuera de rango, como por ejemplo el modelo MHIP [77, 78], descrito brevemente con anterioridad en la Sec. 2.1.1, o LRIP [100, 101]. Estos tres paradigmas han sido comparados [85] desde tres puntos de vista diferentes: según la estructura matemática sobre la que se han construido, según la relación de cada modelo con las leyes físicas o psico-físicas que los justifican, y finalmente, en función del coste computacional de su desarrollo.

El estudio muestra que los tres modelos poseen una base matemática de álgebra lineal. En ese aspecto, tanto MHIP como LRIP, poseen estructuras algebraicas basadas únicamente en espacios vectoriales, con ordenación mediante la estructura de norma. Por otra parte, LIP está basado en una estructura algebraica de *cono lineal con topología ordenada positiva* (ver CONO LINEAL POSITIVO al pie de página), al cual se le proporciona el opuesto, obteniéndose un espacio vectorial completo. Esta estructura algebraica tan robusta proporciona un significado matemático a los conceptos físicos usados en las imágenes. Por tanto, se pueden definir los operadores siguiendo el sentido de una estructura de orden, que desde un punto de vista estrictamente algebraico produce operadores mucho más robustos que los obtenidos siguiendo una estructura de norma, tal y como ocurre en los modelos MHIP y LRIP.

En el citado artículo [85], se demuestra que LRIP no sigue de forma rigurosa ninguna ley física o psico-física; y aunque MHIP sí está fundado en varias leyes físicas y/o psico-físicas, cumple un menor número de leyes que LIP. En la sección 2.2.4 se muestra un resumen de las principales leyes físicas y psico-físicas que cumple el paradigma LIP. Por lo tanto, LIP se puede afirmar que es el modelo más general de los tres, puesto que está definido matemáticamente siguiendo estructuras más generales, permite su aplicación no sólo a imágenes sino a cualquier función matemática y cumple una mayor cantidad de leyes físicas y psico-físicas de construcción de imágenes.

Finalmente, con respecto al aspecto computacional, tanto MHIP como LRIP son dos modelos que trabajan utilizando “transformadas” (es decir, los algoritmos aplican inicialmente una función de transformación sobre la imagen, tras lo que realizan un procesamiento clásico sobre la imagen transformada y finalmente aplican la transformación inversa sobre la imagen resultante). Sin embargo, LIP es un modelo que permite una filosofía de trabajo tanto de tipo “directo” (los algoritmos clásicos se pueden programar usando operadores modificados

LRIP: Del inglés, LOG-RATIO IMAGE PROCESSING. procesamiento de imágenes de razón logarítmica.

Cono Lineal	<i>Se dice que un subconjunto C de un espacio vectorial V es un cono lineal, si y solo si $\lambda x \in C$, $\forall x \in C, \forall \lambda \in V$.</i>
Cono Lineal con Topología	<i>Se dice que C es un cono lineal con topología, si C es un cono lineal y si se cumple que $a \geq b$ entonces $a + c \geq b + c$, y también, $c + a \geq c + b$, $\forall a, b, c \in C$.</i>
Cono Lineal Positivo	<i>Sea C un cono lineal con topología. Se puede obtener el cono lineal con topología ordenada positiva, denominado de manera abreviada, cono lineal positivo $C^+ \subseteq C$, tomando el conjunto de valores $a \in C$ que cumplen que $a \geq 0$.</i>

que proporcionan directamente una imagen resultante, sin necesidad de aplicar ninguna transformada) como de tipo “transformada”, lo que hace que muchos algoritmos se puedan implementar utilizando, al menos, dos mecanismos diferentes.

2.2.4 Leyes físicas y psico-físicas que cumple LIP

Como se ha indicado con anterioridad, además de tener una base algebraica robusta, LIP no es una mera invención matemática, sino que tiene un sustento en la forma en la que el sistema visual humano percibe las escenas. En el modelo LIP, el brillo de un punto de una imagen se ha demostrado que es la cantidad de luz que pasa por un filtro lumínico con una determinada función de absorción. Se ha demostrado [85, 86] que LIP cumple una serie de leyes físicas y de leyes psico-físicas ampliamente aceptadas dentro de la comunidad científica, entre las que cabe destacar las siguientes:

Inversión de escala de tonos de gris

Los límites del intervalo $(0, M)$ representan, respectivamente, el “umbral superior” (también conocido como *límite de deslumbramiento*) y el “umbral inferior” (o *límite de oscuridad total*). El *límite de deslumbramiento* se corresponde con el valor superior de iluminación que el ojo humano es capaz de soportar y distinguir, debido a la intensidad máxima con la que responden los fotorreceptores de la retina humana. Cualquier incremento de la intensidad lumínica por encima de dicho valor no se traduce en un aumento de la respuesta de los fotorreceptores, porque dichos fotorreceptores se encuentran funcionando al límite máximo. Por el contrario, al alcanzar el *límite de oscuridad total*, el sistema visual humano mostraría una ausencia completa de respuesta. Sin embargo, este umbral de oscuridad es un límite teórico inalcanzable, puesto que está demostrado [7, 119] que el ojo humano muestra sensibilidad a una baja cantidad de fotones. Esto se puede probar cerrando los ojos en una habitación oscura y notando como “aparecen” pequeños destellos en la imagen observada, similares al ruido gaussiano. Es decir, que el sistema visual humano proporciona respuesta ante una presencia ínfima de fotones, mostrando que incluso ante una oscuridad plena, el sistema visual humano tiene cierto nivel de sensibilidad, demostrando empíricamente que el *límite de oscuridad total* no es alcanzable en condiciones naturales por el sistema visual humano.

La inversión de escala se justifica de dos maneras. La primera, en el ámbito de los procesos de formación de imágenes por luz transmitida, donde las *funciones de tonos de gris* se deben tomar como filtros de transparencia, de manera que el valor 0 se ha de tomar como *transparencia total*, disminuyendo el grado de transparencia conforme aumenta el valor, hasta tomar el valor de *opacidad total* (M). En segundo lugar, también tiene validez desde el punto de vista de la percepción visual humana. De hecho, se ha demostrado a través de experimentos psicofísicos [7] utilizando monos (que poseen un sistema visual con muchas semejanzas con el sistema visual humano) que en una ausencia completa de estímulos luminosos (simulando una teórica “oscuridad plena”), la intensidad bio-eléctrica de la retina es constante a un valor no nu-

lo y que el aumento de la iluminación en la recepción se traduce en una reducción de dicho nivel de intensidad bio-eléctrica y no en un aumento como cabría esperar. Por tanto, para garantizar la inversión de escala, LIP hace uso de *funciones de tono de gris* en lugar de imágenes en el rango habitual. Gracias a esta inversión, el modelo LIP trabaja de manera similar y consistente a como lo hace el sistema visual humano.

Relación de los tonos de gris con la Intensidad Lumínica

En el contexto de la percepción visual humana, una función de tono de gris \hat{f} se relaciona con el valor de la función de intensidad de la luz incidente F mediante:

$$\hat{f} = M \left(1 - \frac{F}{F_{\text{máx}}} \right) \quad (2.18)$$

donde $F_{\text{máx}}$ es el *umbral superior* o *límite de deslumbramiento* del sistema visual humano, anteriormente citado.

Utilizando (2.18), LIP permite relacionar un concepto físico, la función de intensidad de la luz incidente, F , con una definición propia del funcionamiento de LIP, como es la *función de tono de gris*, \hat{f} , que se corresponde con una función de intensidad de luz incidente valuada en el rango del intervalo real $(0, F_{\text{máx}}]$, donde $F_{\text{máx}} < M$. De hecho, la definición de un *tono de gris* en el contexto de la percepción humana de la intensidad lumínica se puede considerar que es una función normalizada con inversión de escala.

Saturación Lumínica

Al contrario que otros paradigmas, que no están limitados en el rango de intensidades, LIP propone un rango limitado $[0, M]$, que es consistente con el principio de saturación lumínica del sistema visual humano, en el cual, más allá de un cierto límite ($F_{\text{máx}} < M$), el ojo humano no es capaz de reconocer ningún aumento posterior en la cantidad de intensidad de luz incidente.

Ley de Weber

Desde mediados del siglo XIX se sabe que la respuesta del sistema visual humano a la intensidad lumínica no es lineal. Fue Weber el que expuso una ley en la que defendía que el sistema de detección visual humana depende del cociente de la división de los valores de intensidad lumínica más que de la diferencia entre los diversos valores de intensidad. De hecho, introdujo el concepto de "*diferencia de iluminación mínima distinguible*", que es la cantidad de luz necesaria que hay que sumar a un campo de prueba con valor lumínico, F , para que sea visualmente apreciable la diferencia con otro campo de prueba fijo a un valor lumínico F . La ley de Weber se describe:

$$\frac{\Delta F}{F} = W \quad (2.19)$$

donde W es una constante, llamada *Constante de Weber*.

Está demostrado [86] que la operación resta LIP es consistente con la ley de Weber. Aunque esta ley ha sido muy criticada debido a que sólo se cumple para valores de intensidad de luz superiores a un cierto nivel, y a que la constante de Weber sólo es válida dependiendo del tamaño del elemento a detectar. A pesar de todo ello, el modelo LIP es plenamente válido dentro del campo de modelización de la percepción visual humana, ya que en todas las situaciones donde sea cierta la ley de Weber, la resta LIP expresa correcta y coherentemente dicha ley.

Ley de Fechner

Unos años después que Weber, Fechner explicó la no-linealidad del sistema de percepción visual humano de la siguiente manera:

Para producir pasos aritméticos incrementales en la sensación visual, la intensidad lumínica habrá de crecer geoméricamente.

Así pues, introdujo una relación entre la intensidad lumínica en un punto, F , tomado como estímulo de entrada y el brillo observado en dicho punto, B , que representa la sensación visual percibida por el ojo humano en dicho punto. Esta relación dió lugar a la *Ley Discreta de Fechner*:

$$\Delta B = k \frac{\Delta F}{F} \quad (2.20)$$

donde ΔF es el incremento de luz que produce un incremento de la sensación de brillo (ΔB) y siendo k una constante. La *Ley Continua de Fechner* se puede expresar como:

$$B = k' \ln \left(\frac{F}{F_{\min}} \right) \quad (2.21)$$

donde k' es una constante y F_{\min} es el *umbral de funcionamiento absoluto* del ojo humano que representa el valor de iluminación mínima a partir del cual comienzan a distinguirse cambios de iluminación y que se sabe que es muy cercano a la descripción física de oscuridad absoluta. La ley de Fechner se puede escribir también como:

$$B = k' \ln \left(\frac{F}{F_{\max}} \right) + k' \ln \left(\frac{F_{\max}}{F_{\min}} \right) \quad (2.22)$$

donde F_{\max} es el *umbral superior o límite de deslumbramiento* del sistema visual humano, anteriormente citado.

La resta LIP se muestra consistente con la *Ley Discreta de Fechner* y con la *Ley Continua de Fechner*, puesto que aplicando el isomorfismo φ del modelo LIP sobre (2.22) se obtiene (2.23), que demuestra la relación.

$$B = -\frac{k'}{M} \varphi(\hat{f}) + k' \ln \left(\frac{F_{\max}}{F_{\min}} \right) \quad (2.23)$$

De hecho, la ley de Fechner fue uno de los primeros intentos de encontrar una escala lineal para el brillo o sensación de intensidad lumínica con las operaciones clásicas de suma y multiplicación, mientras que LIP define operaciones específicas que trabajan directamente con la función de intensidad lumínica (estímulo de entrada).

2.3 MODELOS DE RUIDO EN IMÁGENES

La tecnología de captación de imágenes por medios digitales está en un continuo avance. Los sensores admiten una mayor resolución y tienen un comportamiento cada vez más realístico. Sin embargo, la captación de imágenes no es precisa y hay una cierta probabilidad de introducir ruido en la imagen digitalizada. Este ruido se puede definir como la alteración apreciable en el valor de un píxel digital determinado con respecto al valor de intensidad asociado a la posición de dicho punto en la imagen real. Se ha estudiado ampliamente el ruido en las imágenes y se han propuesto diversos mecanismos para eliminar completamente (en los casos que sea posible) o, al menos, mitigar el efecto de dicho ruido.

2.3.1 Ruido en imágenes

Inicialmente, se debe precisar qué es lo que se entiende por ruido dentro del ámbito del PROCESAMIENTO DE IMÁGENES. En Bovik *et al.* [12], el ruido se define como *una componente no deseada de la imagen digitalizada*. Sin embargo, esta definición es demasiado amplia y se puede concretar más indicando que:

Se define el *ruido* en las imágenes como toda alteración no deseada en la imagen digitalizada que hace que ésta sea diferente de la imagen real original.

La mayoría de las veces dichas alteraciones son apreciables para el observador, pero otras son pequeñas modificaciones del valor digitalizado de los píxeles que apenas son detectables para el sistema visual humano. Desde el punto de vista matemático, las alteraciones sufridas por un píxel en particular pueden provenir de una modificación del valor del píxel "real" elevando o disminuyendo su valor linealmente. Así pues, una imagen con ruido se puede modelar [42] como:

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y), \quad (2.24)$$

donde $f(x, y)$ es la imagen original, $g(x, y)$ es la imagen original alterada con ruido, $h(x, y)$ es una función de degradación que se aplica espacialmente mediante el operador convolución y $\eta(x, y)$ es un cierto ruido *aditivo* no dependiente de la imagen original. La función $h(x, y)$ también es llamada ruido *multiplicativo*.

Existen una variedad de modelos de ruido, la mayoría de ellos de tipo aditivo. A continuación se expondrán brevemente dos de los más habituales.

RUIDO GAUSSIANO El ruido aditivo más habitual es el gaussiano. Se utiliza para modelar el ruido termal y, bajo algunas ligeras restricciones asumibles, el ruido por conteo de fotones y el ruido por la granularidad de la película fotográfica. Este tipo de ruido se basa en que la probabilidad de que un píxel se modifique se reduce siguiendo una

distribución de probabilidad gaussiana conforme aumenta la diferencia entre el valor de la variación que sufre dicho píxel y el valor real del mismo. Es decir, la probabilidad de que un píxel no sufra variación alguna es elevada, y la probabilidad de que un píxel tenga una variación extrema es muy poco probable. La función de densidad que determina la probabilidad de ruido gaussiano univariado (con media μ y varianza σ^2) es:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.25)$$

donde x puede tomar cualquier valor del intervalo $(-\infty, \infty)$. Como esta definición puede provocar la obtención de valores negativos, aspecto no admisible en imágenes de intensidad, en la práctica el rango de valores de ruido gaussiano se limita a aproximadamente $\pm 3\sigma$, truncando aquellos valores que el ruido gaussiano pudiese llevar a obtener valores negativos [12].

El ruido gaussiano es muy utilizado para modelar el comportamiento del ruido en las imágenes debido al propio proceso de captación de las mismas. Las imágenes son captadas por sensores CCD que miden la energía termal producida por el choque de los electrones incidentes en cada elemento del sensor. Los sensores CCD suponen que la mayoría de los electrones que captan provienen de la luz incidente, sin embargo, existen una cantidad de electrones provenientes del propio material del CCD que, por excitación termal, pueden incidir en el valor captado. Resumiendo, el ruido termal cumple tres características: es el resultado de la vibración de una gran cantidad de electrones, la vibración de cada electrón es independiente de la de los demás y, finalmente, no existen electrones que provoquen una contribución significativa mayor que la de los demás. Estas tres condiciones (fácilmente asumibles) hacen que este ruido se pueda modelar como una distribución de probabilidad Gaussiana, gracias al Teorema del Límite Central [41]. Es decir, el valor captado por cada sensor CCD se puede desviar ligeramente del valor real debido a la presencia de esos electrones excitados por vibración termal; es muy probable que la variación sea nula o muy cercana a cero y es muy poco probable que dicha variación sea muy grande.

CCD: Del inglés, CHARGE-COUPLED DEVICE, dispositivo de carga acoplada.

El ruido de "sal y pimienta" presenta puntos blancos y negros de manera dispersa.

RUIDO DE "SAL Y PIMIENTA" Este tipo de ruido produce una degradación muy característica en la imagen: sólo hay unos pocos píxeles ruidosos, pero la variación (el ruido) que presentan es extrema (tanto hacia el blanco como hacia el negro). Este tipo de ruido aparece típicamente en la transmisión de imágenes por canales digitales con muchas interferencias. En estos casos, cada bit se transmite de manera independiente a los demás, por lo que el cambio de valor de uno de los bits más significativos provoca que cambie completamente su valor debido a una mala decodificación.

2.3.2 Emborronamiento de imágenes

Por la propia naturaleza del ruido, en una primera aproximación se determina que éste se concentra en puntos más o menos aislados, en los que se produce un "salto" en los valores de la intensidad con respecto

a los valores de los píxeles cercanos. Esto hace que el ruido modifique el espectro de la imagen en las altas frecuencias. Debido a esta característica, la primera solución que se suele plantear ante la presencia de ruido es la aplicación de un operador de filtrado paso baja, para eliminar las componentes de altas frecuencias presentes en el ruido. Este filtro no debería ser demasiado severo, puesto que podría eliminar completamente la estructura de la imagen, ya que, como se ha comentado anteriormente, la *reflectancia*, que contiene la estructura de los objetos, también se concentra principalmente en las altas frecuencias. En general, al aplicar un filtro paso baja se produce un cierto emborronamiento de las imágenes, mediante el cual se suele eliminar o mitigar el efecto local del ruido en la imagen, difuminando el efecto entre los píxeles de un entorno.

Para realizar el filtrado paso baja en el dominio espacial es habitual la utilización del operador convolución con máscaras más o menos extensas. Uno de los filtros paso baja más simples es el *filtrado por la media* de entorno $X \times Y$, es decir, el valor de un píxel viene dado por la media de los valores del entorno de dicho píxel (incluyendo o no al píxel en consideración).

Existe otro conjunto de filtros paso baja más complejos, que son los basados en la distribución de probabilidad estadística del ruido. El ruido es una alteración en la medida de un valor y está ampliamente aceptado que la variación en la medición de un valor sigue una distribución estadística de tipo Gaussiana. También se ha estudiado [17] que un píxel no es una variable estadística independiente y que no todo el entorno de un píxel afecta por igual al valor del punto en consideración. Los píxeles más cercanos afectan o influyen más que los más lejanos. Si se supone que el ruido sigue una distribución estadística de tipo Gaussiana, el filtrado de la imagen con una máscara Gaussiana permite que este filtrado afecte de manera controlada a los bordes reales de los objetos. El control de esta operación viene dado por la varianza de la distribución Gaussiana, que permite tener un filtro más amplio, más estrecho, etc., lo que permite modificar el conjunto de píxeles al que afecta la máscara Gaussiana. A esta operación se le suele denominar *filtrado Gaussiano* de varianza σ .

2.4 EXTRACCIÓN DE BORDES

Una de las operaciones fundamentales del procesamiento de imágenes es la detección de bordes, por lo que es un procedimiento muy común en el procesamiento de imágenes estáticas. Los bordes suelen coincidir con los límites de los objetos presentes en las imágenes, aunque también pueden presentarse bordes internos a los objetos, en el fondo, o justo lo contrario, que no se muestre el borde de un objeto porque el cambio de la intensidad de la imagen en el borde no sea abrupto, sino muy gradual, y por lo tanto no se detecte. De manera genérica, existe una gran correlación entre los bordes físicos de los objetos y los bordes representados en las imágenes. Sin embargo, existen casos en los que se pueden marcar como bordes en las imágenes píxeles que no corresponden con ningún objeto físico, como es el caso de las sombras proyectadas.

De las tres grandes familias de operadores que extraen los bordes en las imágenes, Operadores basados en el Gradiente, Operadores basados en el Laplaciano y Operadores Morfológicos, a continuación sólo se introducen las dos primeras, ya que los Operadores Morfológicos no están relacionados con el trabajo desarrollado en esta Tesis.

2.4.1 Filtros basados en el gradiente de la imagen

De manera genérica y, por simplificación, aplicado sobre imágenes en tonos de gris, se define un borde como la localización espacial de un cambio abrupto en valores de tonos de gris espacialmente cercanos. Esta afirmación se puede extender indicando que un *borde se define como la localización espacial donde se da un máximo o un mínimo en la derivada de la función tono de gris, también llamado GRADIENTE de la imagen*. Por simplificación lo que se busca son los máximos en la función valor absoluto del gradiente de la imagen, es decir:

$$|\nabla f(x, y)| \geq T, \quad (2.26)$$

donde T representa el umbral de detección.

Una vez que se obtiene la imagen formada por aquellos píxeles cuyo gradiente es mayor que el umbral T , se suele aplicar un proceso de disminución o adelgazamiento de los máximos detectados, puesto que los extremos detectados suelen estar rodeados de valores altos que superan el umbral establecido y sólo interesan los valores máximos locales de cada entorno detectado en la dirección del gradiente.

En general, las aproximaciones al cálculo de la función gradiente en un espacio discreto, como es el de las imágenes digitales, se suele realizar mediante un par de filtros orientados ortogonalmente, $h_1(x, y)$ y $h_2(x, y)$, que se convolucionan de manera independiente sobre la imagen, de manera que la suma del resultado de ambos es la aproximación a la imagen gradiente:

$$\widehat{\nabla} f(x, y) = (f(x, y) * h_1(x, y)) \cdot \mathbf{u}_1 + (f(x, y) * h_2(x, y)) \cdot \mathbf{u}_2 \quad (2.27)$$

donde \mathbf{u}_1 es un vector unitario en la dirección del filtro h_1 y \mathbf{u}_2 es un vector unitario en la dirección del filtro h_2 .

DETECTORES SIMPLES DE BORDES HORIZONTALES Y VERTICALES

A veces se desea extraer únicamente bordes horizontales o verticales. Para ello, se convolucionan la imagen con filtros que responden sólo en dichas direcciones. Estos filtros pueden usar diferencias entre píxeles a distancia uno o diferencias centrales, que respectivamente son:

$$h_1(x, y) = \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} \quad h_2(x, y) = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \quad (2.28)$$

$$h_1(x, y) = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad h_2(x, y) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \quad (2.29)$$

Se ha resaltado la posición central (x, y) sobre la que se aplica el filtro en la imagen.

Todos los filtros que se propongan han de sumar 0 en sus componentes, ya que implementan una derivada y no deben dar una respuesta ante una constante.

FILTRO DE ROBERTS PARA BORDES DIAGONALES El filtro para la detección de bordes de Roberts [93] está indicado para la detección de bordes diagonales:

$$h_1(x, y) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad h_2(x, y) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (2.30)$$

Este operador presenta un par de problemas. El primero es que el punto de cruce de la diagonal $[-1 \ 1]$ se localiza entre píxeles, es decir, que no coincide con ninguna posición real en la estructura en rejilla de los píxeles. Sin embargo, como es necesario asignar dicho punto de cruce a alguna localización, se producen errores de aproximación generalizados. El segundo problema que presenta este operador (y en general todos aquellos que usan sólo información de dos píxeles) es que su sensibilidad al ruido es bastante alta.

FILTRO DE PREWITT Una posible solución para evitar el primero de los problemas que presenta el filtro anterior es diseñar los filtros en base a las diferencias centrales, para lo cual se utiliza habitualmente un entorno de 3×3 píxeles. Para evitar el segundo de los problemas se suele aplicar un suavizado, que lo que hace es incorporar más píxeles en el proceso de cómputo del valor del píxel central, dando el resultado en función de los píxeles de su entorno, no sólo de 2 píxeles. Teniendo en cuenta esto, para obtener los filtros de Prewitt [87] se definen dos funciones, $h_x(x)$ y $h_y(y)$, que dependen, respectivamente, sólo de x e y . Mediante el producto escalar de estas dos funciones se obtienen unos filtros separables, ortogonales, y que además aplican suavizado al resultado:

$$h_x(x) = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}' \quad h_y(y) = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad (2.31)$$

$$h_1(x, y) = h_x(x) \cdot h_y(y) \\ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad (2.32)$$

Así pues, los filtros del operador de Prewitt son:

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (2.33)$$

FILTRO DE SOBEL Propuesto por primera vez por Irwin Sobel [103, 104] en una charla dentro del Proyecto Stanford Artificial en 1968. Este operador se ha convertido en uno de los más utilizados tanto por la buena respuesta que ofrece como por su relativamente bajo coste

computacional. Es una variante del operador de Prewitt, que utiliza $[1\ 2\ 1]$ en lugar de $[1\ 1\ 1]$, realizando un mejor suavizado. Por tanto, los filtros del operador de Sobel son:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2.34)$$

FILTRO DE FREI-CHEN El filtro de Prewitt tiene más sensibilidad a la detección de bordes horizontales y verticales que a los bordes diagonales, mientras que el filtro de Sobel tiene más sensibilidad a los bordes diagonales que a los horizontales y verticales. Esto es debido a que ambos filtros se han desarrollado sin tener en cuenta la compensación debida a la distancia, que es distinta entre los píxeles diagonales y los píxeles horizontales y verticales. Es decir, que tanto *el filtro de Prewitt* como *el filtro de Sobel* presentan *anisotropía*.

El filtro del operador de Frei-Chen [39] es:

$$\begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix} \quad (2.35)$$

Este filtro, aunque no es totalmente isotrópico, puesto que no es rotacionalmente simétrico, introduce una anisotropía reducida y aceptable.

FILTRO DE SCHARR Existen muchos procesos que exigen filtros con muy alta isotropía, es decir, que se comporten de manera similar sin presentar mayor sensibilidad en la dirección de los patrones evaluados. Scharr et al. [96] presentan un filtro que tiene un comportamiento invariante y optimizado ante la rotación.

$$\frac{1}{32} \cdot \begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix} \quad \frac{1}{32} \cdot \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix} \quad (2.36)$$

2.4.2 Filtros basados en el Laplaciano de la imagen

El Laplaciano de una imagen $f(x, y)$ se define matemáticamente como:

$$\nabla^2 f(x, y) = \nabla \cdot \nabla f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \quad (2.37)$$

Al aplicar la segunda derivada, los bordes de la imagen producirán ceros en esta función, por lo que una de las ventajas que proporciona es que produce bordes sin grosor, puesto que es el cero en sí el que proporciona un único punto de cruce, de manera que no es necesario aplicar un paso posterior de adelgazamiento de bordes. La función Laplaciana (en su versión de función continua) es isotrópica, puesto que no favorece ninguna dirección en particular. Además, los bordes que presenta son contornos cerrados, ya que, al no tener en cuenta

la fuerza de los extremos, sino sólo el paso por cero, el más pequeño cambio gradual produce un paso por cero.

Sin embargo, el Laplaciano presenta algunos problemas:

- Produce la aparición de bordes falsos al detectarse pasos por cero en cambios de convexidad en la función de la imagen.
- Presenta una alta sensibilidad a la presencia de errores, debido al uso de la segunda derivada, creando bordes falsos, al producirse cambios en regiones constantes, o modificando los bordes reales existentes.

El operador Laplaciano es un escalar, frente al operador gradiente que es un vector; por lo tanto, en lugar de un par de filtros ortogonales, es necesario un único filtro para obtener el Laplaciano.

$$\hat{\nabla}^2 f(x, y) = f(x, y) * h(x, y) \quad (2.38)$$

Para evitar el problema de la detección de los pasos por cero entre píxeles, en lugar de usar sólo 2 píxeles, se suele usar un entorno de 3 píxeles, de tal manera que se puede obtener un operador Laplaciano de manera simple de la siguiente forma:

$$\frac{\partial f(x, y)}{\partial x} \rightarrow f_x(x, y) = f(x + 1, y) - f(x, y) \quad (2.39)$$

$$\frac{\partial^2 f(x, y)}{\partial x^2} \rightarrow f_{xx}(x, y) = f(x, y) - f(x - 1, y) \quad (2.40)$$

$$\begin{aligned} \frac{\partial^2 f(x, y)}{\partial x^2} &\rightarrow f_{xx}(x, y) = \\ &= f(x + 1, y) - 2f(x, y) + f(x - 1, y) = \\ &= \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \end{aligned} \quad (2.41)$$

Actuando de manera análoga con y se obtiene:

$$\frac{\partial^2 f(x, y)}{\partial y^2} \rightarrow f_{yy}(x, y) = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad (2.42)$$

Al unificar (2.38), (2.41) y (2.42) (estas dos últimas extendidas a un tamaño de filtro de 3×3), se puede calcular:

$$\hat{\nabla}^2 f(x, y) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (2.43)$$

Mediante otras aproximaciones a la segunda derivada, es decir, proponiendo otras ecuaciones (2.39) y (2.40) y utilizando este mismo método, se pueden construir otros filtros Laplacianos alternativos. Dos ejemplos de filtros Laplacianos distintos del anterior son:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix} \quad (2.44)$$

Laplaciano de Gaussianas (LoG)

Aunque en una imagen pueden existir muchos bordes, el ojo humano es capaz de realizar segmentaciones automáticas y extraer sólo aquellos bordes que le interesan en función de la tarea que se está realizando; para detectar los objetos principales sólo se tienen en cuenta aquellos bordes más abruptos, mientras que para observar la textura de un objeto hay que tener en cuenta todos sus bordes y pliegues. Esta característica ha sido altamente deseada por los científicos con el objetivo de obtener un detector de bordes ajustable según la sensibilidad escogida.

El operador Laplaciano de Gaussianas [67], LoG, también conocido como operador de *Marr-Hildreth*, es un operador que incorpora el concepto de control de sensibilidad. Se ha observado que la convolución de una imagen con un filtro Gaussiano produce un suavizado que limita la imagen resultante a un rango específico de frecuencias, amortiguando el impacto de la presencia de ruido o de bordes débiles (y por tanto no deseados). Si sobre esta imagen filtrada se aplica el operador Laplaciano, se obtienen los pasos por cero en función del grado de sensibilidad que se desee, que es un parámetro controlable por la desviación típica (σ) de la Gaussiana aplicada como filtro.

La Gaussiana presenta unas propiedades que facilitan su uso como filtro del detector de bordes:

- La Gaussiana es una función suave y claramente localizada tanto en el dominio espacial como en el dominio de las frecuencias. Esto permite un comportamiento de suavizado ante los errores a la vez que de precisión en la localización de los bordes reales.
- La Gaussiana es separable, es decir, puede ser aplicada en cada dimensión de manera independiente, lo cual facilita una implementación eficiente.

El filtro Gaussiano en su forma continua se expresa como:

$$g(x, y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.45)$$

Donde el parámetro σ está inversamente relacionado con la frecuencia de corte.

Puesto que tanto el Laplaciano como la convolución son operadores lineales, aplicar un filtrado Gaussiano seguido de una derivación es equivalente a filtrar utilizando la derivada de una Gaussiana, tal y como se muestra en (2.46).

$$\nabla^2 [f(x, y) * g(x, y)] = [\nabla^2 g(x, y)] * f(x, y) \quad (2.46)$$

Al poder calcularse el Laplaciano de una Gaussiana de manera independiente de la imagen sobre la que se aplique, se puede crear previamente un conjunto de filtros para distintos valores de σ , utilizando la ecuación:

$$h(x, y) = \nabla^2 g(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.47)$$

El uso del operador de *Marr–Hildreth* se justifica porque la función que representa (2.47) posee un perfil muy similar al de la respuesta del campo espacial de visión biológica: una respuesta excitatoria circular simétrica rodeada de una banda de inhibición que se termina atenuando. La ecuación anterior se presenta en su forma continua, pero para crear un filtro discreto hay que muestrear (2.47), usando un valor de σ fijado de antemano. Dicho filtro es el que se utilizará posteriormente para convolucionar la imagen. Además, hay que escoger un tamaño suficientemente ancho para que no se produzcan efectos de truncamiento. Una regla habitualmente utilizada es fijar la anchura del filtro a un valor al menos 3 veces mayor que la anchura del pico central de excitación de la Gaussiana. Usualmente, el filtro resultante no suele ser pequeño, con lo cual es más eficiente realizar los cálculos en el dominio de las frecuencias, multiplicando las transformadas discretas de Fourier del filtro y de la imagen, y posteriormente realizarle la transformada inversa al resultado. Para acelerar el proceso de aplicación del filtro espacial, y puesto que la Gaussiana es una función separable, se aplica una convolución del filtro 1-D primero por filas y luego por columnas.

2.4.3 Método de Canny para la extracción de bordes

El detector de *Canny* [17, 43] utiliza tanto la primera derivada (*gradiente*) como la segunda derivada (*Laplaciano*) de las imágenes de una manera muy efectiva y suele ser considerado el detector de bordes más eficaz, aunque su complejidad y coste computacional es elevado. Está diseñado para cumplir los siguientes objetivos:

- **BAJA TASA DE ERROR DE DETECCIÓN.** Todos los contornos reales deben ser detectados.
- **ALTA PRECISIÓN ESPACIAL DE LOS BORDES.** Los contornos detectados han de localizarse tan próximos como sea posible de la posición real.
- **RESPUESTA ÚNICA.** Un borde real debe pertenecer a un único contorno detectado.

Los pasos del algoritmo propuesto por *Canny* [17] son:

A. Obtención del gradiente

1. Suavizar la imagen, aplicando un filtrado Gaussiano (2.45), de manera análoga a como se ha comentado para el Laplaciano de Gaussianas, pero de manera independiente: convolucionar la imagen sobre el eje X con un filtro gaussiano y volver a convolucionar el resultado obtenido sobre el eje Y.
2. Aplicar sobre la imagen suavizada anteriormente una nueva convolución utilizando como filtro la derivada de (2.45) con respecto a los ejes X e Y, respectivamente, obteniéndose dos imágenes convolucionadas.

3. Sobre las dos imágenes doblemente convolucionadas previamente, calcular la magnitud del gradiente y la dirección del mismo, utilizando las siguientes ecuaciones:

$$\left| \widehat{\nabla}f(x, y) \right| = \sqrt{f_{\partial_x}^2(x, y) + f_{\partial_y}^2(x, y)} \quad (2.48)$$

$$\angle \widehat{\nabla}f(x, y) = \arctan \left(\frac{f_{\partial_y}(x, y)}{f_{\partial_x}(x, y)} \right) \quad (2.49)$$

donde $f_{\partial_x}(x, y)$ es el punto (x, y) de la imagen suavizada y derivada en X , $f_{\partial_y}(x, y)$ es el punto (x, y) de la imagen suavizada y derivada en Y , y $\left| \widehat{\nabla}f(x, y) \right|$ y $\angle \widehat{\nabla}f(x, y)$ son, respectivamente, la magnitud del gradiente en el punto (x, y) y el ángulo de dicho gradiente.

- B. *Eliminar puntos del gradiente que no sean máximos locales.* Se determina que todo punto del gradiente es máximo local si su valor es mayor que el de los puntos vecinos en la dirección perpendicular al contorno (es decir, en la dirección del gradiente). Debido a la discretización del espacio, la magnitud del gradiente del punto se compara con el valor promedio del gradiente de aquellos puntos vecinos que están situados en la dirección del gradiente.
- C. *Umbralización mediante histéresis.* Tras los pasos previos hay que aplicar una **BINARIZACIÓN** en la cual se indica qué puntos pertenecen al contorno. Habitualmente, para esta tarea se ha utilizado un único umbral, sin embargo, se ha observado que esta opción no proporciona buenos resultados. Si el valor de un punto está cercano al valor umbral, es probable que otros puntos del contorno cercanos al mismo tengan un valor del gradiente menor y que, por lo tanto, se eliminen del contorno. Para evitar esto, se introducen dos umbrales: el de valor superior identifica aquellos puntos del contorno para los que no hay duda, puesto que tienen un valor magnitud del gradiente bastante elevado, mientras que el umbral de valor inferior determina el valor mínimo por debajo del cual los puntos no son identificados como pertenecientes a ningún borde. Para aquellos puntos del gradiente que tengan un valor entre el umbral inferior y el superior, si están conectados con otro punto que pertenece al contorno, se identificarán como del contorno.
- D. *Síntesis de características de distintos niveles.* Si se aplican sobre una misma imagen distintos filtros gaussianos en los que lo único que se ha variado ha sido el parámetro σ , puede ocurrir que la localización de los bordes difiera un poco entre las distintas imágenes suavizadas. Aunque *a priori* este comportamiento pueda parecer poco robusto, en la práctica es muy útil para la eliminación de los falsos contornos debidos al ruido, ya que es altamente improbable que un error aparezca en diversas imágenes suavizadas. Basándose en estos resultados, **Canny** propuso un mecanismo llamado de “síntesis de características” mediante el cual se unifican los distintos mapas de contornos obtenidos aplicando diferentes

BINARIZACIÓN: Se define como una umbralización de una imagen para obtener una representación en blanco y negro de la misma.

niveles de suavizado mediante un filtro Gaussiano. Esta unificación se realiza siguiendo una filosofía de “fino a grueso”, localizando los contornos en un nivel más fino y buscando un posible pequeño desplazamiento en el nivel inmediatamente más grueso, para ir refinando la localización de dichos puntos del borde real.

Habitualmente este método no se presenta implementado en su totalidad. De hecho, el anterior paso **D** no se suele realizar, puesto que los resultados que se obtienen utilizando el algoritmo hasta este paso son de suficiente calidad y este paso es computacionalmente costoso.

2.5 CONCLUSIONES

En este capítulo se han descrito las características de bajo nivel más relevantes en el ámbito del procesamiento de imágenes. En particular, se ha descrito el modelo multiplicativo de formación de imágenes, que permite conceptualizar de manera más simple los elementos constituyentes que forman las imágenes. Gracias a este modelo, se puede eliminar la componente de *iluminación* y trabajar únicamente con la componente de *reflectancia* o *transmitancia*.

Existen otros modelos de procesamiento de imágenes basados en la respuesta logarítmica de los modelos de formación de imágenes. En concreto, en este capítulo se ha descrito el modelo LIP. El cual cumple una gran cantidad de leyes físicas y psico-físicas, a la vez que está basado en una estructura algebraica robusta, lo que lo hace matemáticamente muy versátil. Este modelo sirve de base para el desarrollo de los operadores especiales que se proponen en esta Tesis Doctoral.

También se ha descrito qué se considera ruido en imágenes y se han expuesto los principales modelos. Otra parte importante de este capítulo se ha destinado a la descripción de los métodos de extracción de bordes. Se han expuesto diferentes mecanismos para la obtención de los contornos, basados en dos operadores matemáticos, el gradiente y el laplaciano. También, se ha mostrado la base matemática que subyace bajo todos estos conceptos elementales y operadores aplicados. Ésta se utilizará en capítulos posteriores para la construcción de nuevos operadores de imagen que permitan la obtención de contornos de objetos en escenas naturales con sombras o con iluminación no uniforme.

Los métodos de extracción de bordes son capaces de producir imágenes en las que se muestran los bordes relevantes. Evaluando la calidad de las imágenes generadas utilizando estos mecanismos de extracción de bordes se podrá escoger el mejor método. En el siguiente capítulo, se describirán diferentes mecanismos para la evaluación de la calidad de las imágenes procesadas.

EVALUACIÓN DE LA CALIDAD DE IMÁGENES PROCESADAS

ÍNDICE

3.1	Evaluación de resultados	46	
3.1.1	Tipos de evaluación de resultados	47	
3.1.2	Evaluación objetiva	48	
3.1.3	Evaluación subjetiva	49	
3.1.4	Restricciones generales	50	
3.2	Contornos de los objetos en imágenes	51	
3.2.1	Importancia de los contornos en imágenes	51	
3.2.2	Conceptos de los contornos en imágenes	51	
3.3	Evaluación de la calidad de contornos	52	
3.3.1	Evaluación objetiva de los contornos de imágenes	52	
3.3.2	Evaluación subjetiva de los contornos de imágenes	54	
3.4	Mean Opinion Score: MOS	56	
3.4.1	Cálculo del MOS	58	
3.4.2	Extensión del MOS	58	
3.5	Metodología de encuestas de imágenes de contornos	60	
3.6	Conclusiones	62	

Un aspecto esencial en el mundo científico es la evaluación tanto de los algoritmos y métodos propuestos por los investigadores para resolver los problemas como los resultados que estos generan. Este aspecto es general a todos los campos de aplicación, por lo que inicialmente se expondrá toda la terminología y casuística general, para luego concretar en el campo del procesamiento de imágenes.

En muchos ámbitos, la evaluación de los métodos o algoritmos se puede realizar comparando los resultados que se obtienen con los esperados, siendo mejor cuanto menor sea la diferencia entre ambos. Esta metodología se puede aplicar si está disponible el resultado esperado y si todas las diferencias que se puedan encontrar entre la respuesta obtenida y la respuesta esperada afectan de igual manera a la bondad del resultados. En el procesamiento de imágenes, la mayoría de las veces no es posible cumplir ninguna de las dos premisas anteriores, por lo que esta metodología no es aplicable directamente.

Existen muchas experiencias que han mostrado que el ser humano es capaz de evaluar la calidad de las imágenes que se le presentan, coincidiendo las respuestas de muchos de los evaluadores. En este capítulo se mostrarán algunas aproximaciones para evaluar de manera subjetiva la calidad de las imágenes procesadas, así como algunos de los problemas y limitaciones que presentan.

3.1 EVALUACIÓN DE RESULTADOS

Todos los ámbitos científicos requieren algún tipo de mecanismo que permita evaluar la bondad de los resultados que proporcionan los métodos diseñados. Esta tarea se denomina **EVALUACIÓN**. En muchos campos del saber, la evaluación se puede reducir a comparar el resultado obtenido con el resultado esperado, considerando cualquier diferencia entre ambos resultados como un error del método evaluado. Dependiendo de cada aplicación, será aceptable una pequeña diferencia en el resultado, es decir, se podrá admitir una cierta cantidad de error en el resultado generado, mientras que en otras aplicaciones, no se aceptará error alguno. Para todas las aplicaciones en las que se pueda aceptar cierto grado de error, se podrá evaluar la **CALIDAD** de los resultados, en función de la cantidad de error cometido. En cualquier caso, para evaluar los resultados de un método se requieren algunas premisas:

- El sistema debe tener, al menos, un resultado “correcto”.
- Debe ser posible proporcionar dichos resultados “correctos” mediante algún mecanismo numérico.
- Los resultados deben estar expresados en un tipo de dato que permita su comparación.
- La comparación de resultados debe proporcionar un valor numérico de la diferencia entre ellos.
- Se tiene que conocer el efecto en la calidad del resultado de la variación tanto por exceso como por defecto del resultado generado frente al valor del resultado “correcto”.

En los casos en los que todas las premisas anteriores se cumplen, la evaluación de los resultados es una tarea sencilla. Sin embargo, la evaluación es muy compleja en todos aquellos sistemas en los que no se cumplen. Por ejemplo, en muchos ámbitos no existe un resultado que sea totalmente “correcto”. En otros sistemas, sólo es posible indicar que los resultados no son iguales entre sí, pero no es posible expresar numéricamente el grado de variación entre ellos.

La última premisa del listado anterior involucra al menos dos métodos, el primero (el método evaluado) que proporciona un resultado utilizado por un segundo método. En estos casos, una mala respuesta del primer método suele empeorar la calidad de respuesta del segundo método. Sin embargo, el error producido puede ser por no alcanzar el valor “correcto” (denominado, *error por defecto*) o por sobrepasarlo (*error por exceso*). Más aún, es probable que una respuesta errónea *por defecto* en el primer método afecte a la calidad de respuesta del segundo método de diferente manera que si el error fuese *por exceso*. Esta situación hace más compleja la evaluación de los resultados.

La **UIT**, ha estado tratando de evaluar la calidad de las señales tanto de audio como de vídeo. Para que las evaluaciones de diferentes sistemas sean comparables, la UIT ha propuesto una serie de documentos, agrupados bajo el epígrafe de **Recomendaciones de la UIT-T** [53, 54]. Éstas plantean la evaluación de la calidad de los resultados de cualquier método, en función de sus salidas, desde dos puntos de vista:

VALORACIÓN Clasificación absoluta de la calidad intrínseca del método propuesto. Dentro de las Recomendaciones de la UIT-T se denomina *Absolute Category Rating*, ACR.

COMPARACIÓN Clasificación relativa de la calidad del método propuesto con respecto a otros. Estos otros pueden ser métodos diferentes ampliamente aceptados, o bien, el mismo método con diferentes parámetros de ejecución. En relación con el primero de los casos, la UIT-T compara pares de métodos entre sí y lo denomina *Pair Comparison*, PC. En cuanto al segundo caso, se conoce como *Degradation Category Rating*, DCR dentro de las Recomendaciones de la UIT-T.

Mediante la **VALORACIÓN** de un método se determina la calidad del resultado obtenido. Para ello, es imprescindible conocer *a priori* el resultado "correcto" y disponer de una métrica capaz de medir la diferencia entre éste y el resultado generado por el método propuesto.

Por otra parte, la **COMPARACIÓN** permite evaluar la calidad relativa de varios métodos en función de sus respectivos resultados. Para ello, es necesario que la respuesta de todos los métodos se encuentre normalizada (dentro del mismo rango) de forma que se les pueda aplicar una métrica que permita compararlos (y ordenarlos) en función de las salidas. Esta medida no garantiza que el método que proporcione el mejor resultado comparativo posea un resultado aceptable, simplemente determina qué método es el mejor entre la batería de métodos comparados.

3.1.1 Tipos de evaluación de resultados

Una primera clasificación de los tipos de evaluación permitiría diferenciar los métodos de evaluación en dos grandes grupos:

OBJETIVOS La calidad de un resultado se proporciona mediante un valor numérico generado utilizando algún tipo de algoritmo, función o método matemático. También son llamados métodos **CUANTITATIVOS**, **ANALÍTICOS** o **EMPÍRICOS**.

SUBJETIVOS La calidad de un resultado no es proporcionada mediante un valor numérico, sino que es valorada por seres vivos, habitualmente humanos. Se les denomina también métodos **CUALITATIVOS**.

Los métodos objetivos para la evaluación de los resultados son más rápidos que los subjetivos, ya que permiten la automatización de los cálculos mediante la aplicación sistemática de los algoritmos de evaluación. Para aquellos sistemas y métodos que cumplan todas las premisas descritas con anterioridad se podrá obtener un método objetivo

UIT: UNIÓN INTERNACIONAL DE TELECOMUNICACIONES. También conocida por sus siglas en inglés, ITU

UIT-T: UNIÓN INTERNACIONAL DE TELECOMUNICACIONES, sector Normalización de TELECOMUNICACIONES

para su evaluación. En los sistemas en los que alguna de las premisas no se cumpla, también es posible determinar de manera objetiva la calidad de un resultado mediante la utilización de un segundo método que utilice el resultado que se desea evaluar y que produzca resultados evaluables objetivamente. Se les denomina *EVALUACIÓN OBJETIVA INDIRECTA*. En estos casos, se estima que la calidad de la respuesta de este segundo método dependerá directamente y en exclusiva de la calidad de los resultados del primer método.

Por otra parte, los métodos subjetivos requieren la evaluación de los resultados por seres vivos, lo que hace que su aplicación sea, en general, lentos. Habitualmente, estos métodos se utilizan cuando se desea evaluar resultados directamente utilizables por humanos, por ejemplo, la calidad de un equipo de sonido o la calidad de imagen de televisión. En general, muchos de los sistemas evaluados mediante métodos subjetivos podrían ser evaluados con métodos objetivos indirectos. Sin embargo, dicha evaluación puede llegar a ser excesivamente compleja, ya que pueden existir parámetros que tengan una gran influencia en las respuestas de los métodos finales y para los que el ajuste óptimo sea tan complicado como la evaluación en sí. Además, estas medidas objetivas indirectas no garantizan que el resultado evaluado tenga mayor calidad en cualquier otro objetivo diferente del que presentaba el segundo método.

3.1.2 Evaluación objetiva

Existen muchos métodos objetivos, entre los más utilizados cabe destacar *MSE* y *PSNR*. Estos métodos pertenecen al grupo de métodos denominados *MÉTRICAS DE DATOS* [117]. Están compuestos por fórmulas sencillas de comprobar y de calcular. En general, en un amplio espectro de aplicaciones muestran un comportamiento coherente con respecto a la calidad. Sin embargo, estos métodos están basados en comparaciones byte a byte de las muestras, sin tener en cuenta qué representa cada elemento. Tampoco se tienen en cuenta posibles relaciones adicionales entre los mismos. Por ejemplo, en la evaluación de imágenes, los bytes pertenecientes a los píxeles no tienen un tratamiento especial y tampoco se tienen en cuenta las relaciones espaciales de cercanía entre los píxeles. Esto es así, porque estos métodos fueron diseñados para caracterizar la fidelidad de los datos entre dos copias, sin tener en cuenta el contenido, ni el significado ni las relaciones que los datos puedan mantener entre sí.

Para cada campo de aplicación se han diseñado métricas específicas, por ejemplo, para el ámbito del procesamiento de imágenes existen otras aproximaciones, que se engloban dentro de las llamadas *MÉTRICAS DE IMAGEN* [117], que proporcionan medidas de la calidad de las imágenes. Algunos de los métodos de evaluación incorporan componentes del sistema visual humano, como la percepción del color, sensibilidad local del contraste, búsqueda de patrones específicos, etc. utilizando modelos y datos que han sido obtenidos de experimentos psicofísicos relacionados. A este tipo de técnicas se les denomina *APROXIMACIONES DEL MODELO VISUAL*.

Dentro de las *MÉTRICAS DE IMAGEN* existen otros métodos agrupados

MSE: Del inglés, MEAN SQUARED ERROR, error cuadrático medio
PSNR: Del inglés, PEAK SIGNAL-TO-NOISE RATIO, relación pico señal-ruido

bajo el nombre de APROXIMACIONES DE INGENIERÍA. En este grupo de métodos se encuentran todos aquellos que extraen y analizan determinadas características o buscan la presencia de ciertas distorsiones presentes en las imágenes de menor calidad. En caso de encontrarlos, calculan la fuerza con la que se presentan dichas características o distorsiones y proporcionan una medida de calidad global. Estas técnicas no rechazan completamente los conceptos del grupo de APROXIMACIONES DEL MODELO VISUAL, sino que hacen uso de él para determinar como se generan dichas distorsiones y buscarlas.

En general, todos los métodos objetivos se pueden agrupar en tres categorías [18] en función de los requerimientos que plantean para su funcionamiento:

REFERENCIA COMPLETA Métricas que requieren tanto la respuesta “correcta” como la respuesta obtenida por el método evaluado.

REFERENCIA REDUCIDA Métricas que requieren la respuesta obtenida por el método evaluado y una descripción parametrizada de la respuesta “correcta”.

SIN REFERENCIA Métricas que únicamente requieren la respuesta obtenida por el método evaluado.

Las críticas que se hacen a todos los métodos OBJETIVOS es que presentan poca flexibilidad para ajustarse a otros ámbitos de actuación, incluso aunque sean relativamente “cercaños” al que han sido diseñados. Además, en muchos campos de investigación no existen teorías sólidas que permitan trasladar con total exactitud las opiniones o comportamientos humanos a fórmulas matemáticas y, por lo tanto, cualquier intento de obtener una métrica objetiva no puede ser más que una aproximación, sin garantía en la precisión del resultado con respecto a la calidad percibida por un humano.

3.1.3 Evaluación subjetiva

Todos aquellos sistemas que generen respuestas que puedan ser tratadas por seres humanos, podrán ser evaluados por estos. Este tipo de evaluación se denomina EVALUACIÓN SUBJETIVA. Los valores de la evaluación subjetiva pueden variar mucho de un humano a otro, por lo que es necesario realizar la evaluación utilizando muchas personas como evaluadores. También es habitual utilizar muchas entradas de prueba suficientemente similares y representativas para reducir el sesgo todo lo posible.

En algunos ámbitos de investigación, la evaluación subjetiva es el único mecanismo posible para determinar la calidad de los métodos. Por tanto, este tipo de evaluación ha sido utilizada ampliamente y existen muchos estudios realizados en este aspecto. Es predominante la utilización de la evaluación subjetiva en investigaciones de ciencias sociales, ciencias de la salud y humanidades, y en particular, en estudios que analizan la calidad de servicios.

La evaluación subjetiva presenta tres partes diferenciadas:

PREGUNTAR Realizar las preguntas pertinentes a los evaluadores para recabar su opinión sobre la calidad de los resultados de los métodos bajo evaluación.

RECOGER RESPUESTAS Captar las opiniones de los evaluadores y traducir las respuestas cualitativas de los evaluadores a valores cuantificables.

PROCESAR RESPUESTAS Computar todas las respuestas, valorarlas y proporcionar una respuesta unificada representativa de todas las respuestas proporcionadas por los evaluadores.

Como se indicará en secciones posteriores, los resultados proporcionados por los métodos diseñados en esta Tesis Doctoral se evaluarán utilizando métodos subjetivos. Por consiguiente, las tres partes que componen la evaluación subjetiva serán descritas con mayor profundidad en posteriores secciones y capítulos de esta Tesis.

3.1.4 *Restricciones generales*

Para algunos campos de aplicación, el resultado “correcto”, necesario para el proceso de VALORACIÓN, es conocido y fácilmente calculable. Sin embargo, en otros campos de aplicación, el resultado “correcto” no es conocible *a priori*, el proceso de cálculo necesario para su obtención es complejo o, simplemente, el cálculo del valor exacto es inabordable computacionalmente, por lo que no se podrá aplicar una valoración objetiva del método. Por otro lado, aunque se disponga del resultado “correcto”, si el resultado del método evaluado no es el mismo, su variación con respecto al resultado “correcto” se puede tratar de manera diferente en cada campo de aplicación e incluso en cada aplicación particular. Por ejemplo, para una aplicación determinada, puede ser admisible que el resultado del método propuesto difiera en un 5% del resultado “correcto”, mientras que para otra aplicación puede que esta diferencia sea inadmisibile. Todo ello hace que las evaluaciones comparativas sean más comunes que las valoraciones objetivas.

Para poder comparar y ordenar los resultados es necesario que exista una relación de orden entre ellos. Sin embargo, en algunos casos dicha relación de orden no es fácil de determinar o simplemente es dependiente de cada aplicación concreta. Es decir, que en ciertas aplicaciones no es posible proponer un mecanismo que permita indicar que un resultado es mejor que otro. Esto puede deberse a que, para un sistema en particular, no se ha conseguido especificar el comportamiento ante errores por exceso o por defecto en el resultado generado. Para proponer una relación de orden que permita comparar y ordenar los resultados, se hace necesario restringir el ámbito de la evaluación. Esto permite aplicar conocimientos de tipo heurístico gracias a lo cual se puede indicar cómo varía la calidad ante errores por exceso o defecto.

Como puede deducirse de la exposición anterior, la evaluación de la calidad de los resultados, estimada como la aproximación a un resultado “correcto”, es una de las tareas principales en la mayoría de investigaciones científicas y, además en muchos casos, con un alto grado de complejidad. Tratar de realizar una taxonomía completa de todos

los métodos de evaluación de la calidad en todos los campos de aplicación se escapa del ámbito del presente trabajo. Así que a partir de ahora, esta Tesis se ceñirá al campo del procesamiento de imágenes, y en particular, a la extracción de bordes o contornos en imágenes.

3.2 CONTORNOS DE LOS OBJETOS EN IMÁGENES

Dentro de las aplicaciones relacionadas con el campo del procesamiento de imágenes, una de las primeras tareas que se han realizado desde el comienzo de esta disciplina ha sido la obtención de los bordes relevantes de los objetos presentes en las mismas. En esta sección, se mostrará la importancia que presentan los bordes o contornos de los objetos para el reconocimiento de las escenas, los conceptos tanto matemáticos como psicológicos asociados a dichos bordes y las problemáticas asociadas a la detección de los mismos.

3.2.1 Importancia de los contornos en imágenes

Para entender la importancia que presentan los bordes en la visión humana, [Cohen \[22\]](#) realizó un experimento en el que utilizaba un espacio de color uniforme y sin ningún objeto en él. Este experimento mostró que cuando el ojo humano observa un espacio uniforme y sin bordes durante un periodo de tiempo elevado, independientemente del color que pueda tener dicho campo, tiende a transformar dicho espacio en un campo con un color uniforme de tonalidad gris oscura, denominado **GANZFELD**. En dicho experimento también se constató que tan pronto se introducía cualquier objeto o cambio de iluminación, dicho *Ganzfeld* desaparecía y los sujetos eran capaces de volver a ver el espacio en su color uniforme junto con el nuevo objeto introducido. [Coren, Ward y Enns](#) exponen [24] que un efecto similar aparece de manera natural al permanecer en un entorno con mucho hielo y nieve durante un tiempo prolongado, en estos casos muchos sujetos afirman tener una sensación de “no poder ver”, a lo que se ha llamado *ceguera de la nieve*. Esto demuestra la necesidad de que existan bordes en una imagen para que el sistema visual humano sea capaz de actuar. [Matlin y Foley \[69\]](#) afirman que “mediante esta experiencia se demuestra empíricamente un principio básico de la visión humana: sin bordes, no hay visión”. Por tanto, de ahí la importancia de los bordes en las imágenes, y por consiguiente, de la necesidad de detectarlos para poder realizar posteriores procesamientos.

GANZFELD: Del alemán, “campo total”.

3.2.2 Conceptos de los contornos en imágenes

Antes de trabajar con los bordes, es necesario especificar qué se considera borde o contorno de los objetos en el ámbito del procesamiento de imágenes. Así pues, desde una perspectiva geométrica [12], un BORDE o CONTORNO se puede definir de dos formas diferentes, bien como el conjunto de puntos en los que se produce un cambio abrupto en la orientación apreciable de una superficie física, o bien como el conjunto de puntos que separan dos o más regiones diferentes de uno o más

elementos físicos. Por otro lado, y desde un punto de vista perceptual [69], bajo una perspectiva física–matemática, un BORDE se podría definir como la posición espacial donde existe un CAMBIO REPENTINO en el *brillo*, *luminosidad* o *color* de la imagen. Donde la *luminosidad* se define como el reflejo acromático (sin color) percibido de la superficie (ej: diferentes tonos de blancos, negros, sombras de diversos tonos de gris, etc.). Por otra parte, el término *brillo* representa la cantidad percibida de energía luminosa incidente en un cierto punto de la imagen (ej. un foco luminoso puede ser brillante o tenue, pero no podrá ser negro, blanco o gris).

Sin embargo, ninguna de las definiciones anteriores contempla el concepto psico–físico de un borde *relevante*. El problema de dicha definición es que es por naturaleza DEPENDIENTE DEL CONTEXTO y además, en muchos casos, DEPENDIENTE DEL OBSERVADOR. Es decir, que para una misma imagen, dos aplicaciones diferentes, podrían encontrar dos conjuntos de bordes *relevantes* distintos, e incluso dos observadores podrían no coincidir los bordes seleccionados.

Es más, la *relevancia* de los bordes, no es un concepto dicotómico, sino que es un concepto con diversos grados de *relevancia* (al estilo de la *Lógica Difusa*). Es decir, no es correcto afirmar que un determinado borde es *relevante* (o, al contrario, no es *relevante*), sino que su grado de *relevancia* es elevado (o, al contrario, su grado de *relevancia* es bajo). Por tanto, no es posible proponer una definición única de qué es un borde *relevante*, puesto que es necesario un mecanismo que permita determinar el grado de *relevancia* del mismo.

3.3 EVALUACIÓN DE LA CALIDAD DE CONTORNOS

A continuación, se exponen las características asociadas a los métodos de evaluación de la calidad de bordes. No se detallará ningún método en particular, sino que primero se describirán las características inherentes a los métodos objetivos. Tras lo cual, se incidirá en las características de los métodos subjetivos.

3.3.1 Evaluación objetiva de los contornos de imágenes

La evaluación del rendimiento de los algoritmos de extracción de bordes ha estado presente desde los comienzos de esta disciplina, bien para poder determinar la precisión de los bordes extraídos, o bien para comparar el resultado de nuevos algoritmos propuestos con otros ya existentes. Como se ha comentado al inicio del capítulo, en muchos campos, la evaluación de los resultados es una tarea sencilla, ya que basta con comparar el resultado obtenido por el método con el objetivo (conocido *a priori* y que no posee errores). Sin embargo, en el campo de las imágenes en general, y en el de la extracción de bordes en particular, esto no es tan simple. En la mayoría de los casos, no es posible proporcionar un resultado “correcto”, aunque se pueden diseñar experimentos “controlados” con imágenes para las cuales se conoce de antemano el resultado esperado. Normalmente, estos experimentos están compuestos por IMÁGENES SINTÉTICAS, o por imágenes reales bajo

unas condiciones muy estrictas, por lo que no son útiles para simular entornos reales, ya que no proporcionan información visual real.

De manera mayoritaria, los mecanismos de evaluación de los bordes han sido de tipo cuantitativo, es decir, sin incluir la intervención humana. Para ello se han propuesto una gran cantidad de métricas [38, 118], muchas de ellas basadas en el uso de IMÁGENES SINTÉTICAS, comentadas con anterioridad, o bien en el uso de GROUND TRUTH, obtenido a través del consenso en el resultado de varios métodos [37] o mediante *segmentación manual de uno o varios expertos* [51, 72]. Las principales ventajas que presentan estos métodos son que, al no requerir de la intervención humana, son muy rápidos, se pueden aplicar en batería a muchas imágenes resultado, y proporcionan un valor numérico fácilmente comparable. Sin embargo, presentan dos problemas mayores: el primero es la dificultad (o incluso, imposibilidad) para la obtención del GROUND TRUTH y el segundo, que los métodos objetivos están diseñados con una idea prefijada de lo que se considera el mejor resultado; este hecho implica que si dicha idea cambia, será necesario rehacer la métrica completamente.

GROUND TRUTH:
Imágenes resultado
con el valor correcto

Por otra parte, al existir diferentes algoritmos para la extracción de los bordes de una imagen, cada uno de ellos puede producir imágenes en las que los bordes pueden variar espacialmente, no ser detectados en algunos casos, identificar como bordes posiciones que no lo son, etc. En este contexto, el principal problema reside en cómo medir el error ante la presencia o ausencia de un borde, suponiendo que se conoce el resultado "correcto".

Una posible solución al problema de la presencia o ausencia de bordes podría ser el uso de la evaluación cuantitativa, obtenida a partir de 4 valores calculados comparando la imagen resultado con la imagen resultado objetivo [88]:

- Bordes existentes en la imagen original que han sido detectados en la imagen resultado (*Positivos Correctos*).
- Bordes inexistentes en la imagen original que NO han sido detectados en la imagen resultado (*Negativos Correctos*).
- Bordes existentes en la imagen que NO han sido detectados en la imagen resultado. (*Falsos Negativos* o *Bordes No Detectados*).
- Bordes inexistentes en la imagen original que han sido detectados en la imagen resultado. (*Falsos Positivos* o *Falsas Alarmas*).

La estimación del error está relacionada con la cantidad de *Falsos Negativos* y de *Falsos Positivos* que tenga una imagen resultado con respecto a la imagen resultado objetivo.

De manera adicional, otro aspecto complejo es la determinación del resultado "correcto", ya que éste debe contener únicamente todos los bordes *relevantes* de la imagen, y como se ha expuesto con anterioridad, el grado de *relevancia* de cada borde dependerá de muchos factores, la mayoría de ellos de carácter subjetivo.

3.3.2 Evaluación subjetiva de los contornos de imágenes

Antes se ha indicado que la imagen resultado se comparaba con la imagen resultado objetivo; sin embargo, dicha imagen objetivo no se suele poseer, ya que este objetivo suele depender del campo de aplicación al que se aplique el método: un mismo método en un campo de aplicación puede dar unos resultados excelentes, mientras que aplicado a otro campo de aplicación puede generar unos resultados nefastos. Esto lleva a reiterar la problemática de la subjetividad al evaluar la calidad de las imágenes procesadas: la calidad de los resultados depende del ámbito de aplicación y la evaluación debe ser realizada de manera subjetiva por humanos.

El único mecanismo ampliamente aceptado como válido [15, 71, 116, 117] para la evaluación de la calidad en el ámbito del procesamiento de imágenes y vídeo es mediante test subjetivos. Sin embargo, muchas aplicaciones se evalúan utilizando exclusivamente métodos objetivos, proporcionando una estimación del error en función de la cantidad de aciertos (*Positivos Correctos* y *Negativos Correctos*) frente a errores (*Falsos Negativos* y *Falsos Positivos*). Pero, ¿cómo afecta la presencia de estos errores a la calidad de la imagen resultado? La respuesta a esta pregunta depende de la aplicación para la que se vaya a utilizar la imagen resultado. En algunos casos, la detección de "bordes inexistentes" (debidos al ruido o a la textura) puede provocar que la imagen resultado no permita la identificación de los objetos presentes. Sin embargo, en otros casos, un exceso de bordes no será negativo, ya que proporcionará más datos a niveles superiores, que podrán ser eliminados en fases posteriores. También puede ocurrir que nos encontremos con que los bordes de un determinado objeto no sean detectados completamente, lo que podría llegar a impedir la identificación o la localización de dicho objeto (por ejemplo, en el caso de que falten los trazos principales de una letra). En este caso, la eliminación de un borde es un aspecto altamente negativo. Sin embargo, en otros casos, la no detección de ciertos bordes no es tan nociva, y aunque una parte del objeto no sea detectada, esto no impide su identificación (por ejemplo, que falten algunos trazos no principales de una letra, que no impida su lectura). Por tanto, la estimación del error no va a depender tanto del número de bordes *Falsos Negativos* o *Falsos Positivos* detectados como de qué bordes sean y de cómo afecte su detección o su pérdida a la hora de interpretar el resultado.

En aplicaciones con procesamiento de imágenes cuyo resultado se va a mostrar directamente a un humano, como por ejemplo, la ayuda a la navegación para personas con problemas de BAJA VISIÓN [112], el realce de objetivos para soldados en entornos militares [36], etc., la calidad de los bordes que se muestren al usuario dependerá de un factor muy subjetivo: lo bien o mal que los bordes obtenidos permitan al usuario realizar la tarea que desee en cada momento. Esta dependencia estará relacionada con la disminución en el esfuerzo que la persona haya tenido que realizar para llevar a cabo dicha acción con respecto a si la hubiese desarrollado sin ayuda del soporte artificial. Esto quiere decir que si el procesamiento de bordes produce una salida que ayude poco o incluso que moleste mucho al usuario para realizar una tarea en particular, la calidad del procesamiento de dichos bordes habrá si-

do baja. Por el contrario, si el usuario es incapaz de realizar una tarea por sí mismo, cualquier pequeña ayuda, aunque no sea suficiente para permitir la realización completa de dicha tarea, podría significar que el procesamiento de los bordes ha sido positivo (aunque, ciertamente, no suficiente). Con estos simples ejemplos, se muestra la problemática y la dificultad que tiene la determinación de la calidad de un procesamiento de bordes.

Aunque se puede considerar que muchos métodos objetivos presentan una parte subjetiva, que es la búsqueda del GROUND TRUTH mediante técnicas de anotación manual, en esta Tesis se considerarán como subjetivos aquellos métodos o técnicas en los que la calificación de la calidad de las imágenes de bordes se realice por seres humanos. Es habitual que tras esta evaluación por humanos se aplique algún tipo de procesamiento estadístico posterior, que proporcione un valor numérico comparable.

Un estudio científico [71] avala mediante un análisis estadístico riguroso la necesidad de utilizar la evaluación subjetiva en el ámbito del procesamiento de imágenes, particularizando en los mecanismos de evaluación de algoritmos de segmentación. En dicho estudio, *McKoen et al.* realizaron una evaluación de prueba de un algoritmo de segmentación, obteniendo las siguientes conclusiones sobre el proceso de evaluación:

1. Se deben determinar qué aspectos psico-físicos influyen en las respuestas de los evaluadores para especificar los requisitos subjetivos de las tareas que se pretenden evaluar.
2. Las evaluaciones subjetivas de la calidad de los algoritmos de segmentación permiten contextualizar de manera explícita los diferentes ámbitos de trabajo y, por tanto, permite escoger expertos en cada uno de ellos.
3. Debido a que los objetivos de la aplicación (en este caso, la segmentación de las imágenes) son muy dependientes del contexto de trabajo, es imprescindible una definición concreta de dicho contexto. Esta definición permite determinar las diversas tareas de evaluación subjetiva.

Ejemplos de evaluación subjetiva en imágenes

Para mostrar la importancia que tiene la evaluación subjetiva de la calidad de las imágenes procesadas, se mostrarán algunos ejemplos de organismos y entidades internacionales que han desarrollado trabajos, recomendaciones o que han amparado estudios en este ámbito. Aunque en estos ejemplos no se evalúan de manera específica las imágenes de bordes, su campo de aplicación es muy cercano al de los extractores de bordes. Así pues, gran parte de las propuestas que se realizan en estos ejemplos pueden servir como antecedente y base para el desarrollo de métodos subjetivos para la evaluación de la calidad de imágenes de bordes.

La **UIT** ha desarrollado diferentes recomendaciones para evaluar la calidad de emisión y recepción de sonidos, imágenes y vídeo tras distintas codificaciones y compresiones. Estas recomendaciones cubren

aspectos como el entorno ambiental en el que se deben realizar las evaluaciones, duración de las evaluaciones, número y tipo de las evaluaciones, etc. Muchas de estas recomendaciones pueden exportarse a otras aplicaciones, como la extracción de bordes, segmentación de objetos, etc.

La Unión Europea ha amparado en varios Proyectos Marco el desarrollo de técnicas de segmentación de vídeos. En particular el Grupo de Trabajo **COST 211QUAT** [26], ha trabajado para el desarrollo de técnicas de segmentación de objetos en movimiento en vídeo y la evaluación comparativa (COMPARACIÓN) de los métodos propuestos. El Grupo **COST 211QUAT** solicitó a los investigadores y científicos propuestas de algoritmos de segmentación para compararlos con el método de segmentación propuesto por el propio Grupo. Para evaluar los métodos que enviaban los investigadores y científicos se realizaban una serie de medidas de parámetros objetivos de los métodos y posteriormente una evaluación subjetiva, por parte de unos cuantos usuarios, de unos parámetros previamente fijados. También solicitaban criterios objetivos para la evaluación de la segmentación. Sin embargo, no había planteamientos de como unificar los resultados de la evaluación de los aspectos objetivos y la evaluación de los conceptos subjetivos. Tampoco se presentó ningún mecanismo para darle mayor importancia a unos parámetros (objetivos o subjetivos) frente a otros. A pesar de que esta metodología no es exportable directamente a la evaluación subjetiva de imágenes de bordes, el método seguido puede tomarse como punto de partida y *hoja de ruta* de cualquier propuesta de evaluación subjetiva en imágenes. Por otra parte, gracias a la metodología utilizada en este trabajo, otros autores propusieron [71] un método de evaluación adaptable a cada aplicación, mediante el ajuste previo de los requerimientos, y un análisis estadístico final con el cual determinar las diferencias estadísticamente significativas entre las respuestas de los diferentes métodos.

3.4 MEAN OPINION SCORE: MOS

En 1990, el antiguo **CCIR** (*Comité Consultivo Internacional de Radiocomunicaciones*), precursor de la **UIT**, sentó unas bases mínimas para la evaluación subjetiva de **CÓDECS** digitales de vídeo, publicadas mediante un informe [19]. La **UIT** retomó este trabajo y en 1994 publicó una recomendación [52], en la que se proponía un método para la evaluación subjetiva de la calidad de las imágenes de televisión. Posteriormente, a partir de 1996, la **UIT** publicó una serie de recomendaciones internacionales para la evaluación de la calidad mediante métodos subjetivos, que utiliza una medida basada en la opinión subjetiva de los usuarios. Esta medida ha sido denominada **MOS** [55]. La **UIT** ha ido publicando periódicamente recomendaciones sobre el tema y mejorando las técnicas de experimentación para la evaluación de la calidad, aunque todas ellas siguen basadas en MOS. En aquellos campos de aplicación para los que sea necesaria la utilización de las capacidades de catalogación de los usuarios, por ejemplo, para los casos de segmentación o de extracción de bordes relevantes, es imprescindible utilizar una métrica que permita medir las decisiones subjetivas de cada usuario con

CÓDECS: Abreviatura de Compresor-
Decompresor de
imágenes, audio y
vídeo.

MOS: Del inglés:
MEAN OPINION
SCORE, nota media
de opinión.

VALOR	CALIDAD
5	Excelente
4	Bueno
3	Aceptable
2	Mediocre
1	Mala o Inaceptable

Tabla 3.1
Escala calidad absoluta (ACR) de 5 niveles discretos.

Valor	Calidad
9	Excelente
8	
7	Bueno
6	
5	Aceptable
4	
3	Mediocre
2	
1	Mala o Inaceptable

Tabla 3.2
Escala calidad absoluta (ACR) de 9 niveles discretos.

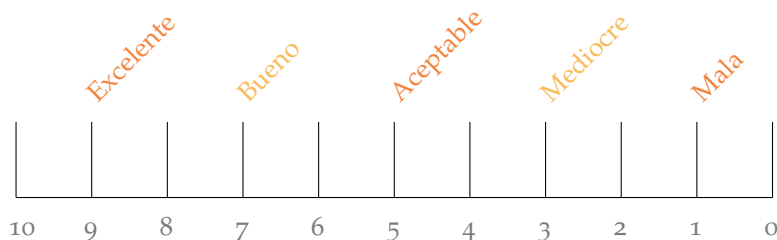


Figura 3.1
Escala calidad absoluta (ACR) de 11 niveles con *puntos de anclaje*.

respecto a un parámetro. En todos estos casos, MOS puede utilizarse como la medida normalizada de las opiniones subjetivas de los usuarios.

Para evaluar la calidad global de cada imagen resultado siguiendo el formato de evaluación de VALORACIÓN (o siguiendo la nomenclatura de la UIT, evaluación ACR) se debe presentar la imagen a cada usuario-evaluador, preguntándole donde situaría la calidad de cada imagen en una escala de enteros valuados entre 1 y 5. Dicha escala tiene la asociación numérico-lingüística mostrada en la Tabla 3.1.

Si se requiere una evaluación más precisa se puede utilizar una escala de 9 u 11 niveles, que puede ser tanto numérica discreta (ver Tabla 3.2) como continua (ver Fig. 3.1). En el caso de la Figura 3.1, se muestra una ampliación de esta escala, en la que los puntos extremos se definen verbalmente como "*puntos de anclaje*". En esta definición verbal se utiliza alguna forma de referencia, que podría ser implícita o explícita, y que debe ser claramente definida durante la fase de instrucción de los usuarios antes de la presentación del cuestionario. En este caso, el valor 10 indicaría una calidad totalmente fiel al original o una calidad de procesamiento perfecta. Frente a esto, el valor 0 indica una calidad

de imagen sin parecido alguno a la original (no es posible imaginar una calidad peor) o un procesamiento totalmente erróneo.

3.4.1 Cálculo del MOS

Tras obtener las respuestas de opinión de cada usuario para cada pregunta y asignarlas a cada algoritmo o método, se calcula la nota media de opinión y los intervalos de confianza (habitualmente al 95 %). En (3.1) se muestra la fórmula para calcular el valor MOS para cada uno de los métodos o algoritmos que se están evaluando.

$$\text{MOS} = \frac{1}{N} \sum_{i=1}^N V_i \quad (3.1)$$

Donde V_i es el valor de cada uno de los votos emitidos y N es el número total de votos.

Para calcular el intervalo de confianza (al 95 %) de los diferentes valores MOS, según cada uno de los diferentes métodos, se aplica (3.2).

$$95\% \text{CI} = \mu \pm \frac{t \cdot \sigma}{\sqrt{N}} \quad (3.2)$$

Donde μ es el valor MOS, t es el estadístico *t-Student* [41] de dos colas para un nivel de significación de $0.05\% = 2.069$, σ es la desviación típica de las medidas entre los diferentes usuarios y N es el número de medias de usuarios.

3.4.2 Extensión del MOS

Inicialmente, *McKoen et al.* [71] propusieron un método estadístico para calcular la diferencia en el rendimiento de distintos algoritmos. Como primer paso se determinó qué parámetros eran subjetivos y cuáles eran objetivos. En una segunda etapa, los encuestados visualizaron las secuencias de vídeo que se deseaban evaluar y finalmente rellenaron una encuesta. Esta encuesta estaba compuesta por 4 preguntas por cada secuencia y donde las respuestas a cada pregunta estaban puntuadas mediante una **ESCALA DE LICKERT** de 5 puntos. Con respecto a los parámetros subjetivos, tras la recogida de todas las respuestas proporcionadas mediante encuestas, los encuestadores realizaron un análisis estadístico ANOVA de la varianza para determinar diferencias estadísticamente significativas entre las medias de las puntuaciones de respuesta. Por otra parte, para los parámetros objetivos, se obtuvieron los resultados de manera directa y se compararon entre sí. En ningún caso, se mezclaron las respuestas de los parámetros subjetivos y objetivos. Esta metodología no permite unificar parámetros objetivos y subjetivos, es decir, no es capaz de proporcionar una respuesta única que englobe todas las respuestas. De manera adicional, otra crítica que se le hace a esta metodología es que todos los parámetros tienen el mismo peso en la respuesta global.

Debido a muchas de las críticas indicadas en el párrafo anterior, *Durucan et al.* [31], ampliaron este método para realizar la evaluación práctica del rendimiento de unos algoritmos de detección de movimiento.

ESCALA DE LICKERT:
Escala psicométrica
comúnmente
utilizada en
cuestionarios, en el
que cada respuesta
se contesta indicando
el grado de acuerdo o
desacuerdo con una
declaración
(elemento, ítem o
reactivo).

En su artículo se presenta un mecanismo para integrar valoraciones objetivas y subjetivas. También se introduce un mecanismo para ponderar las diversas valoraciones, de manera que se pueda dar mayor o menor énfasis *a priori* a las respuestas de algunas preguntas. Esto permite que las diferentes preguntas de un cuestionario puedan tener distinta relevancia, que a su vez permite modular un mismo cuestionario de manera diferente en función del campo de aplicación concreto en el que se estén valorando los distintos métodos y algoritmos.

Así pues, Durucan *et al.* definieron una función de rendimiento para evaluar métodos de segmentación de objetos en movimiento sobre un determinado conjunto de escenas de prueba (diferentes vídeos procesados según el método que se estuviese evaluando). Dicha función se muestra en (3.3).

$$P = \sum_{v=1}^{N_v} p_v \quad (3.3)$$

Donde P muestra el valor de rendimiento completo de un método para todas las N_v secuencias procesadas. A su vez, p_v se encuentra definida en (3.4), y representa el valor de la función de rendimiento para una secuencia en particular, obtenida como la suma del grado de cumplimiento de los N_o requisitos objetivos más los N_s requisitos subjetivos de cada grupo.

$$p_v = \sum_{i=1}^{N_o} p_{oi} + \sum_{j=1}^{N_s} p_{sj} \quad (3.4)$$

Donde todos los requisitos están ponderados por una función de peso, dependiendo de su importancia relativa entre sí, como muestra (3.5),

$$\begin{aligned} p_{oi} &= W_{oi} \cdot G_{oi} \\ p_{sj} &= W_{sj} \cdot G_{sj} \end{aligned} \quad (3.5)$$

Donde p_{oi} representa el valor del requisito *objetivo* i -ésimo, W_{oi} proporciona la ponderación de dicho requisito y G_{oi} indica la respuesta del usuario para el mismo. Análogamente, p_{sj} representa el valor del requisito *subjetivo* j -ésimo, W_{sj} proporciona la ponderación para dicho requisito y G_{sj} indica la respuesta del usuario para el mismo.

Si se asignara la máxima puntuación en cada uno de los apartados de cada una de las secuencias procesadas, se obtendría el valor de rendimiento máximo, es decir, el rendimiento que tendría un “*método perfecto*”. Por tanto, el rendimiento teórico comparado se puede calcular como el porcentaje de la función de rendimiento de cada método con respecto al valor de rendimiento máximo del teórico “*método perfecto*”. Aunque originalmente este método estaba propuesto para el análisis de la calidad de los algoritmos de segmentación, es fácilmente adaptable a la evaluación de imágenes de bordes.

Como se puede deducir, el principal inconveniente de esta técnica es la dificultad que presenta la asignación de ponderaciones a los distintos requisitos. Esto ha limitado su aceptación generalizada dentro del ámbito científico.

3.5 METODOLOGÍA DE ENCUESTAS DE IMÁGENES DE CONTORNOS

La evaluación subjetiva de imágenes de bordes no sólo puede basarse en el uso de una cierta métrica, es necesario utilizar una metodología que permita incorporar, de manera controlada y normativa, la subjetividad dentro de las respuestas de los evaluadores. Aunque tanto *McKoen et al.* como *Durucan et al.* utilizaron encuestas para obtener las respuestas que deseaban, ninguno siguió un método contrastado y, por tanto, no se podría garantizar que sus cuestionarios estuviesen **VALIDADOS** (ver 4.8.1, pág. 82). Además ninguno de sus trabajos estaba específicamente diseñado para trabajar con imágenes de bordes, por lo que, no hay certeza plena de que los métodos que propusieron pudieran funcionar en este ámbito de trabajo. Sin embargo, *Heath et al.* fueron precursores en la catalogación subjetiva de imágenes de bordes y propusieron [47–50] una metodología robusta, en particular, para la búsqueda del conjunto de parámetros que, de manera subjetiva, proporcionase la mejor segmentación de la imagen. En la Tab. 3.3 se especifican los 4 criterios principales para la evaluación del rendimiento de los algoritmos de extracción de bordes, propuestos por dichos autores.

VALIDADOS: Un cuestionario se dice que está **VALIDADO** si hay garantía de que realmente sirve para medir estadísticamente las variables para las que se diseñó.

Tabla 3.3
Criterios generales para la evaluación de la calidad de los algoritmos de extracción de bordes.

- A. La evaluación de los algoritmos de extracción de bordes se debe realizar en el contexto de una tarea de visión concreta.
- B. Se debe determinar el ámbito de trabajo específico (dominio de la tarea) y seleccionar imágenes representativas de dicho dominio.
- C. Se debe utilizar un mecanismo eficiente e imparcial de selección de parámetros que no favorezca a algún algoritmo de extracción de bordes frente a otros y que sea coherente con el ámbito de actuación en el que se desarrolle la evaluación.
- D. La evaluación de los **DETECTORES DE BORDES** debe realizarse sin que los participantes sepan qué algoritmo está siendo evaluado en cada imagen.

La metodología de trabajo propuesta partía de varios algoritmos de extracción de contornos y varias imágenes catalogadas previamente, que representaban una gran variedad de objetos y de contextos, todos ellos fácilmente reconocibles por humanos. Se utilizaron un conjunto amplio de configuraciones para los parámetros, basadas en las propuestas por los autores en los trabajos originales de los algoritmos de extracción de bordes. Partiendo de esas configuraciones, se generaron imágenes de bordes y se evaluaron por un experto que realizó un “pre-pilotaje”, mediante el cual se redujo el grupo de configuraciones a unas pocas, que proporcionaban, a juicio del experto, los mejores resultados de imágenes de bordes. Tras esta primera fase de selección, se generaron diferentes imágenes de contornos utilizando cada una de las diferentes configuraciones escogidas en la etapa anterior. También se escogieron

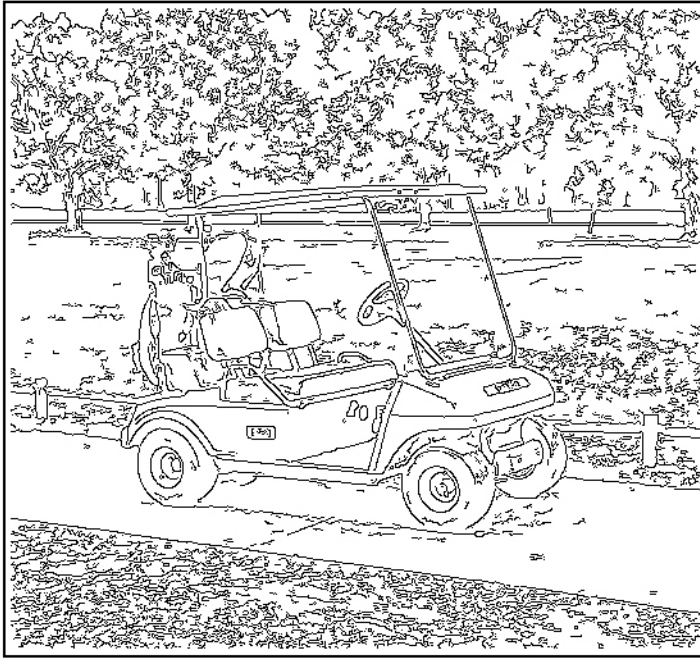


Figura 3.2
ÍTEM del cuestionario
de HEATH *et al.* [49]

Information allows
for easy, quick and
accurate recognition
of the object.

7

6

5

4

3

2

1

No coherent
information from
which to recognize
the object.

los sujetos que posteriormente realizarían las evaluaciones. Todos ellos eran estudiantes voluntarios con alto grado de formación en PROCESAMIENTO DE IMÁGENES. Los autores no proporcionan mayor información con respecto al proceso de selección de la población ni a características específicas de los mismos, por lo que, cabe interpretar que ninguno presentaba problemas de visión.

Finalmente, se pasó a la fase de evaluación mediante cuestionarios. Las imágenes de bordes obtenidas mediante el conjunto reducido de configuraciones para los parámetros de los diferentes algoritmos fueron evaluadas sin límite de tiempo por los individuos anteriormente seleccionados. Se les permitió comparar las distintas imágenes tantas veces como desearan. Los encuestados valoraron cada imagen de contornos respondiendo a la pregunta:

Califique la calidad de los bordes para reconocer el objeto central presente en la imagen original.

La respuesta se encontraba graduada utilizando una *escala de Lickert* entre 1 y 7, donde 1 significaba “Los bordes no proporcionan información coherente que permita reconocer el objeto” y 7, “Los bordes proporcionan información que permite un reconocimiento sencillo, rápido y preciso del objeto”. En la Fig. 3.2 se muestra un ITEM del cuestionario.

3.6 CONCLUSIONES

En este capítulo se han mostrado los dos tipos diferentes de mecanismos de evaluación de resultados, los métodos OBJETIVOS y los métodos SUBJETIVOS. En particular, se ha podido observar que, con carácter general y debido a las características especiales de las imágenes, los métodos de evaluación de éstas han de ser SUBJETIVOS. Para aplicar estos mecanismos, se ha descrito la métrica MOS, junto con la extensión de la misma propuesta por Durucan *et al.*. También se ha descrito la metodología de Heath *et al.* para la evaluación de las imágenes de bordes.

Los trabajos de Heath *et al.* sirven como un punto de inicio contrastado científicamente para la creación de encuestas sobre la calidad subjetiva de imágenes de bordes. Esta metodología, sin embargo, presenta ciertas limitaciones, por ejemplo, no es posible incorporar información de tipo objetivo, todas las preguntas han de tener el mismo grado de importancia en la respuesta global, todas las preguntas son del mismo tipo (todas tienen un mismo texto de pregunta y se utiliza la misma *escala de Lickert*), por lo que no hay posibilidad de realizar preguntas dicotómicas (Sí/No), etc. Gran parte de estas limitaciones se podrían resolver utilizando el método propuesto por Durucan *et al.* utilizando las respuestas obtenidas a partir de encuestas construidas según la metodología propuesta por Heath *et al.*.

Tras este capítulo se puede afirmar que el mecanismo más fiable para evaluar la calidad de las imágenes de bordes es mediante algún mecanismo SUBJETIVO. La evaluación SUBJETIVA requiere conocer la opinión de diferentes usuarios, obtenido utilizando ENCUESTAS o CUESTIONARIOS. En el Capítulo 4 se describirán sus características, los inconvenientes que presentan, los problemas asociados y las recomendaciones que, con carácter general, se han de seguir en su diseño para construir buenos CUESTIONARIOS.

ÍNDICE

4.1	Encuestas de opinión mediante cuestionarios	64
4.2	¿Qué es un cuestionario?	65
4.2.1	Definición de qué es un cuestionario	65
4.2.2	Características buscadas con los cuestionarios	67
4.3	Mecanismos de recolección de las respuestas	67
4.3.1	Entrevistas estructuradas	67
4.3.2	Cuestionarios auto-administrados	68
4.3.3	Tipos de preguntas	69
4.4	Limitaciones de los cuestionarios	70
4.5	Fuentes de error en los cuestionarios	71
4.6	Elección de la muestra poblacional	74
4.6.1	Selección de individuos	74
4.7	Diseño de un cuestionario	76
4.7.1	Dificultades para construir buenas preguntas	76
4.7.2	Buenas preguntas con buenas respuestas	79
4.8	Análisis del cuestionario	82
4.8.1	Validación del cuestionario	82
4.8.2	Fiabilidad del cuestionario	83
4.8.3	Alfa de Cronbach	83
4.8.4	Coficiente de Correlación Intraclase	85
4.9	Conclusiones	87

Habitualmente, la captación y la evaluación de la opinión de usuarios se realiza utilizando cuestionarios o encuestas. Estas encuestas trasladan la opinión de la población (o al menos de una parte representativa de la misma) mediante las preguntas de dichos cuestionarios. La literatura científica actual sobre cuestionarios se centra principalmente en cuestionarios de tipo “lingüístico” en áreas de investigación económicas o sociales. Por tanto, el diseño de cuestionarios adecuados para la evaluación de la calidad del procesamiento de imágenes es un gran reto.

En este capítulo se presentarán algunas ideas de cómo deben diseñarse los cuestionarios y qué problemas y limitaciones presentan los mismos. Se describirán diferentes tipos de sesgo que pueden afectar a las encuestas y se mostrarán preguntas de ejemplo que presentan algunos de los problemas y sesgos descritos. Finalmente, se expondrán brevemente varios métodos de análisis de cuestionarios para estudiar la validez y la fiabilidad de los cuestionarios.

4.1 ENCUESTAS DE OPINIÓN MEDIANTE CUESTIONARIOS

Diversos autores han estudiado las etapas en las que se divide un *estudio de investigación*. Entendiendo dicho *estudio de investigación* como un proceso científico mediante el cual se pretende evaluar la veracidad de unas hipótesis de trabajo a partir de datos obtenidos de una población (sea humana o no). Burgess [16] propone una lista concisa de 7 pasos:

1. Definir el propósito de la investigación; fijar los objetivos de la misma.
2. Identificar la población y la muestra poblacional.
3. Determinar cómo recabar las respuestas.
4. Diseñar el cuestionario.
5. Realizar una encuesta piloto en una muestra poblacional controlada.
6. Realizar la encuesta global con el cuestionario final.
7. Analizar los datos de la encuesta global.

La lista anterior es bastante general y cada etapa puede presentar una complejidad elevada, así pues, Oppenheim [76] amplía y profundiza la lista, extendiendo la misma a 12 etapas o pasos:

1. Decidir cuáles son los *propósitos* del estudio, tanto generales como específicos. Traducir estos *propósitos* en *objetivos funcionales* (esto es, el conjunto de hipótesis que se desean investigar). Mediante estos *objetivos funcionales* se construirá la definición de las *variables* que se desean medir mediante las *preguntas*.
2. Realizar una revisión bibliográfica de la literatura científica.
3. Conceptualizar de manera preliminar el estudio: Entrevistas exploratorias.
4. Decidir el diseño del estudio y determinar su viabilidad.
5. Decidir qué hipótesis se investigarán: crear un listado de *objetivos funcionales* y de variables que se desean “medir”.
6. Determinar la muestra poblacional que será necesaria para obtener un muestreo válido: Representatividad, grupos de control, grupo de seguimiento, etc.
7. Seleccionar los sujetos de la muestra poblacional que participarán el estudio de manera efectiva.
8. Realizar el trabajo de campo: proceso de recolección de datos mediante entrevistas o cuestionarios.
9. Procesar los datos: codificar los resultados y preparar los datos para su posterior análisis.

10. Analizar los resultados mediante un estudio estadístico.
11. Conjuntar todos los resultados y comprobar que las hipótesis se cumplen con los datos obtenidos.
12. Escribir el informe final.

Ambas listas están diseñadas teniendo en mente la finalidad de la obtención de resultados válidos, por lo que es de vital importancia definir correctamente los objetivos que se persiguen con el estudio de investigación. Estos objetivos son los que determinarán las demás etapas, ya que mediante la definición de los objetivos podremos determinar la población destino y qué muestra poblacional debemos tomar de dicha población.

Por ejemplo, si el objetivo deseado es estudiar el grado de aceptación de un nuevo biberón en Andalucía para una posible comercialización del mismo, la población destino serán los bebés menores de 1 año. De esta manera, teniendo en cuenta la concreción espacial y los límites temporales, se reduce el universo completo de seres humanos del mundo: el estudio se centra en los bebés menores de un año que viven en Andalucía. También se puede especificar la muestra poblacional que se tomará para obtener resultados estadísticamente útiles. Para ello, se determina qué individuos se escogerán de entre la población inicial partiendo desde el objetivo inicial: calculando el grado de interés de cada individuo conforme se vayan tomando de la población, o bien, escogiéndolos previamente.

En general, como se puede observar en ambas listas propuestas, el desarrollo de una encuesta y la captación efectiva de las respuestas de los entrevistados son sólo dos de las etapas que componen el *estudio de investigación*.

4.2 ¿QUÉ ES UN CUESTIONARIO?

Los cuestionarios son los mecanismos más habituales para recoger información de las encuestas de opinión. El diseño de los cuestionarios es una de las etapas más importantes dentro de un estudio de investigación mediante encuestas de opinión. Un buen diseño del cuestionario, seleccionando bien las preguntas y las respuestas posibles, permitirá obtener respuestas de los entrevistados con las que proporcionar una conclusión que describa la opinión de dichos entrevistados.

4.2.1 Definición de qué es un cuestionario

La RAE presenta en la 22^a Edición del Diccionario de la Lengua Española [89] una entrada para la palabra *cuestionario* que en su 2^a acepción se define como:

RAE: Real Academia Española de la Lengua

CUESTIONARIO: sust. m. (lat. *quaestionarius*)

2. "Lista de preguntas que se proponen con cualquier fin".

Esta definición, aunque muy formal y académica, es muy vaga, y si no se contextualiza, es poco útil para su aplicación en un ámbito científico, por lo que, un cuestionario se podría definir como:

CUESTIONARIO: *“Conjunto de listas de preguntas que se utilizan para obtener información sobre lo que la gente opina, piensa o siente relativo a un servicio, producto o cuestión en particular, con un propósito específico previamente fijado por el investigador que lo ha diseñado”.*

Estas listas de preguntas pueden estar impresas o no, se pueden rellenar directamente por parte de los entrevistados o mediante entrevistadores que realicen las preguntas y cataloguen las respuestas, pueden realizarse de manera individual o grupal, etc. Como se puede observar, la taxonomía de cuestionarios que se ha presentado es extremadamente amplia, lo cual, aunque parezca un contrasentido, puede llegar a ser tanto una ventaja como una complicación.

La ventaja se presenta porque se tiene disponible una herramienta tremendamente flexible, que va a permitir el acceso a muchas personas desde diferentes frentes de exploración y la obtención de muchos datos, que pueden ser:

DATOS DIRECTOS Respuestas explícitas de los entrevistados. Ej: *¿Qué edad tiene Vd.?*

DATOS INDIRECTOS Deducciones obtenidas a partir de las respuestas de los entrevistados, pero no obtenidas directamente de una pregunta. Ej: Podemos obtener el rango de riqueza (rico, medio, pobre) a partir de los ingresos anuales, de los gastos habituales, etc.

RELACIONES OCULTAS Deducciones obtenidas al cruzar los datos de una población con diversas variables y de las que previamente no se conocía su interrelación. Ej: Deducir que en una población de ancianos en un lugar específico el principal problema deja de ser el de *subir escaleras* conforme los ancianos comienzan a interaccionar diariamente con otros ancianos fuera de su hogar.

La complicación en el diseño de las encuestas viene por la gran versatilidad disponible al existir tantos formatos diferentes de preguntas y tantas modalidades distintas para la captación de las respuestas. Además, *a priori*, no es posible saber si un mecanismo de captación de respuestas es mejor que otro para una determinada población, si una pregunta está bien formulada para obtener un resultado válido, etc.

Los cuestionarios son herramientas muy flexibles y versátiles para obtener la opinión de una población.

4.2.2 Características buscadas con los cuestionarios

El principal cometido que se persigue con un cuestionario es:

Conseguir respuestas de los entrevistados basadas en OPINIONES PROPIAS que proporcionen una INFORMACIÓN GENERALIZABLE mediante la cual contrastar la veracidad de unas hipótesis.

De la aseveración anterior hay que resaltar unos conceptos:

- Se desean obtener respuestas que expresen OPINIONES PROPIAS, por lo que la respuesta de cada entrevistado debe provenir exclusivamente de su propio conocimiento y capacidad de decisión, y en ningún caso debe estar influenciada por elementos externos, es decir, la respuesta no debe estar SESGADA.
- Las respuestas de los entrevistados deben proporcionar INFORMACIÓN GENERALIZABLE. La cual involucra todas aquellas evidencias e indicios que son comunes a un colectivo, es decir, que no son exclusivas de un individuo, por lo que de ellas se podrán deducir opiniones válidas para una población.

Para que las opiniones que se obtengan de los individuos sean UNIVERSALES, la muestra de individuos a los que se les realice la entrevista ha de ser seleccionada minuciosamente en una fase previa al diseño de las preguntas. Esta selección debería garantizar que la muestra poblacional escogida representa completamente a la población universo.

4.3 MECANISMOS DE RECOLECCIÓN DE LAS RESPUESTAS

Un aspecto muy importante, que determina en gran modo el diseño de los cuestionarios, es definir cómo se obtendrán las respuestas de los encuestados. Existen 2 tipos principales de mecanismos de recolección de respuestas:

- A. Entrevistas estructuradas.
- B. Cuestionarios auto-administrados.

A continuación, se describen los dos mecanismos de recolección de datos, tras lo que se expondrán brevemente los tipos de preguntas más habituales.

4.3.1 Entrevistas estructuradas

La recolección de respuestas mediante “entrevistas estructuradas”, involucra a un entrevistador, que guía la entrevista según las respuestas del entrevistado. Estas entrevistas se realizan generalmente cara a cara o mediante llamada telefónica. El hecho de que el entrevistador realice de manera oral las preguntas, con una *realimentación* inmediata por

parte de los entrevistados, hace que las frases se propongan de manera diferente a como se haría si estas preguntas se hicieran por escrito. En este tipo de cuestionarios, además de incluir frases para los entrevistadores, se les debe instruir para que sepan responder a las dudas que les surjan a los entrevistados, así como para que sean capaces de llevar una encuesta dinámica y ágil. Usualmente, estas encuestas son de corta duración (no suelen superar los 10 minutos en el peor de los casos), y están recomendadas para la obtención de estados de opinión puntuales (por ejemplo, en elecciones y votaciones).

Sus grandes ventajas son: su inmediatez y su alta tasa de respuestas (salvo para encuestas que involucren aspectos de tipo sensible, en los que responder directamente a una persona pueda provocar rechazo). En cuanto a sus desventajas el coste de este tipo de encuestas es elevado, lo que limita bastante el número de encuestados. Además, las respuestas tienen una alta probabilidad de ser “políticamente correctas” al hacerse delante de una persona. Por último, su principal problema es que sólo pueden ser de tipo general, es decir, que no involucren procesos mentales excesivamente complejos, ya que en dichos casos, la mayoría de los encuestados abandonan.

4.3.2 *Cuestionarios auto-administrados*

El otro gran grupo de cuestionarios se denomina, de manera general, “cuestionarios auto-administrados”. Al contrario que los anteriores no involucran la presencia activa de un entrevistador durante el proceso de recolección de las respuestas. Por tanto, es necesario añadir unas instrucciones precisas para los entrevistados, y que las preguntas se diseñen de forma que no surjan dudas. Se suele admitir que estos cuestionarios no se responden en una única sesión, así que deben diseñarse para poder dejarlos y continuar en otro momento. Debido a esto, los “cuestionarios auto-administrados” suelen ser más largos e incluir más preguntas que los que se realizan mediante “entrevistas estructuradas”. Este mecanismo está específicamente indicado para poblaciones muy dispersas y difícilmente accesibles mediante entrevistas directas.

Presentan ciertas ventajas: son mucho más baratos que las “entrevistas estructuradas”, permiten recolectar datos de procesos mentalmente más complejos (ya que el encuestado tiene más tiempo para responder) y permiten hacer preguntas de aspectos más sensibles. Sin embargo, entre desventajas destacan su bajo nivel de respuesta, su lentitud y, sobre todo, el poco control que se tiene sobre las respuestas: el entrevistador “tiene que fiarse” de que las respuestas están realmente meditadas (que no han sido respondidas sin reflexionar), y de que han sido proporcionadas por el propio entrevistado (no por otra persona).

4.3.3 Tipos de preguntas

La tipología de preguntas posibles en un cuestionario es muy amplia, a continuación se muestra un breve resumen de las principales categorías:

A. SEGÚN LA ELECCIÓN DE LAS RESPUESTAS

ABIERTAS Todas las respuestas son posibles, de manera que no existe una lista de respuestas entre las que escoger.

CERRADAS Sólo se admiten como posibles respuestas las existentes dentro de una lista de opciones previamente diseñada por el encuestador.

SEMICERRADAS Se le proponen al encuestado una lista de posibles respuestas (habitualmente, las consideradas más comunes) y se le proporciona la opción de añadir otra respuesta de manera libre.

B. SEGÚN EL TIPO DE LAS RESPUESTAS

DICOTÓMICAS Las respuestas son del tipo verdadero/falso, sí/no u otras opciones similares.

CATEGORIZADAS O DE ESCALA ORDINAL Las respuestas suelen ser adjetivos o adverbios, que describen algún tipo de graduación subjetiva, con los que los encuestados responden a las preguntas.

VALORACIÓN O DE ESCALA NUMÉRICA Las respuestas son valores numéricos que tienen asociado un orden que permite a los encuestados proporcionar su grado de aceptación o rechazo respecto a lo planteado en la pregunta. También pueden indicar el orden relativo entre las distintas opciones planteadas.

C. SEGÚN LA CAPTURA DE LOS DATOS DESEADOS

DIRECTAS Si se pregunta directamente al encuestado por la variable que se desea medir.

INDIRECTAS Si se pregunta al encuestado por otra cuestión, que permite determinar la variable deseada.

D. SEGÚN EL NÚMERO DE RESPUESTAS POR PREGUNTA

RESPUESTA ÚNICA Sólo se admite una respuesta por pregunta.

RESPUESTA MÚLTIPLE Es admisible proporcionar varias respuestas para una misma pregunta. Dentro de este tipo de preguntas se tienen dos divisiones:

- *Respuestas múltiples sin repetición:* Se admiten varias respuestas por pregunta pero no es posible repetir varias veces la misma respuesta.
- *Respuestas múltiples con repetición:* Se admiten varias respuestas para una misma pregunta y es posible seleccionar la misma respuesta varias veces.

4.4 LIMITACIONES DE LOS CUESTIONARIOS

Antes de estudiar más a fondo los cuestionarios, conviene saber qué se puede obtener utilizando los estudios basados en encuestas. Estos estudios pueden servir tanto como instrumentos de recogida de datos, como mecanismos para el análisis de variables indirectas. Esta gran versatilidad permite su aplicación en muchas temáticas diferentes. Por tanto, para todos aquellos que se acercan por primera vez a las encuestas existe una tendencia generalizada a pensar que éstas “*son comodines que sirven para todo*” [2]. Sin embargo, las encuestas muestran limitaciones y condicionantes que hacen que éstas no sean válidas en todos los casos.

Por ejemplo, las encuestas no son apropiadas para investigar las razones, motivos o causas subjetivas de comportamientos o fenómenos poco conocidos, en poblaciones pequeñas o poco frecuentes y/o difíciles de acceder. La encuesta es un método dirigido a la descripción y a la constatación de hipótesis, más que al descubrimiento o a la elaboración de hipótesis o teorías. A continuación, se mostrarán algunas de las críticas más habituales que se hacen a las encuestas.

FALTA DE OBJETIVIDAD Los resultados de las encuestas dependen del contexto del significado e interpretación de las preguntas por parte de los entrevistados. Además, la perspectiva teórica que se haya escogido influye en la información recogida (*sólo se puede encontrar lo que se busca*).

INFLUENCIA EN LAS RESPUESTAS Las propias preguntas realizadas en las encuestas pueden crear opiniones *ex novo* en individuos que no han reflexionado nunca sobre un tema.

ATOMISMO La información que surge de la encuesta, en muchos casos, es una mera agregación de respuestas individuales que no tienen en cuenta interacciones sociales, relaciones, ni otros aspectos relativos a la estructura social que no se hayan tenido en cuenta en el modelo conceptual con el que se ha diseñado el cuestionario.

CAPTURA ESTÁTICA DE LA REALIDAD Los resultados de las encuestas muestran, en el mejor de los casos, una “fotografía” estática de las opiniones en un momento puntual. Esto obligaría a realizar estudios en distintos instantes temporales, con la consiguiente dificultad para integrar todas las opiniones.

LIMITACIÓN DEL ÁMBITO DE APLICACIÓN Las encuestas no están previstas para ámbitos de estudio en los que las medidas no se puedan obtener mediante preguntas subjetivas. Por ejemplo, en los ámbitos de ciencias naturales es absurdo realizar los estudios utilizando encuestas, puesto que en la mayoría de los casos, no se puede preguntar a los sujetos (no son humanos, pueden ser animales, plantas, etc.). Además, en estos ámbitos las medidas objetivas son más precisas y se pueden obtener directamente.

COSTES ELEVADOS La realización de un estudio de investigación mediante cuestionarios es caro, tanto en tiempo como en costes económicos y medios materiales y humanos.

Como resumen, se puede afirmar que las encuestas son instrumentos estructurados de captura de información, que podrían influir en la información recogida. Los cuestionarios son útiles, ante todo, para describir y contrastar hipótesis o modelos. Por el contrario, no son recomendables para construir teorías o para deducir hipótesis nuevas.

4.5 FUENTES DE ERROR EN LOS CUESTIONARIOS

La validez de los resultados que se obtengan de un cuestionario dependerá de qué se ha preguntado, cómo se ha preguntado, a quién se le ha preguntado, qué opciones se han ofrecido para responder, en qué contexto se ha realizado la pregunta y otros aspectos relacionados con la encuesta y con su administración. A veces, los resultados obtenidos a partir de una encuesta varían con respecto a la opinión real de la población, obtenida mediante otras encuestas o bien utilizando otros cuestionarios que han sido validados previamente. Dicha variación es debida a la presencia de errores, que pueden ser *sistemáticos* o *aleatorios*. Los errores *aleatorios* no tienen gran incidencia en el conjunto de la encuesta, puesto que ni sobreestiman ni subestiman los valores promedio, y si la muestra poblacional es suficientemente grande, pueden mitigarse fácilmente. Sin embargo, los errores *sistemáticos* pueden llegar a producir un falso aumento o disminución del valor promedio, produciendo un **SESGO** en los resultados. Estos sesgos pueden tener su origen en el diseñador de la encuesta, el encuestador o los entrevistados y, a su vez, pueden ser intencionados o no. La mayoría de errores provienen del diseño del cuestionario y de las preguntas, dentro de las que las principales fuentes de error son [20]:

SESGO: Error que aparece en los resultados que puede conducir a conclusiones que son sistemáticamente diferentes de la verdad o incorrectas.

ERRORES DE ESPECIFICACIÓN CONCEPTUAL El modelo subyacente sobre el que se construye la encuesta está mal construido.

- A. *Problemas de connotación:* Confusión de significados porque el modelo está asociado a más de un significado o porque hay varios modelos diferentes que pueden responder al mismo significado y los investigadores no han hecho explícito lo que pretenden.
- B. *Problemas de denotación:* Modelos débilmente definidos y “vagos” en cuanto a su precisión, en los que no están claramente especificados los objetos, entidades o referentes a los que se aplica.
- C. *Problemas terminológicos:* Fallos de elección de etiquetas lingüísticas o conceptuales para definir el modelo que inducen a características o referentes erróneos.

ERRORES DEBIDOS A LA FORMULACIÓN DE LAS PREGUNTAS Las preguntas o las opciones de respuesta provocan que los encuestados escojan una opción distinta a la opinión real que tienen sobre los objetivos de investigación asociados a cada pregunta.

- A. **SESGO DE ESPECIFICACIÓN:** No existe una correspondencia entre la pregunta y los objetivos de la investigación: se está haciendo una pregunta que no mide realmente lo que se había planteado en los objetivos del cuestionario.

- B. **SESGO DE MEDICIÓN:** La pregunta está bien planteada y está acorde con los objetivos del cuestionario, pero la respuesta que proporcionan los encuestados no coincide con su opinión real.
1. *Sesgo de interpretación de términos:* Se produce al introducir palabras ambiguas o con diferentes connotaciones, que pueden ser interpretados por los encuestados de diferente manera a como fueron ideadas por el diseñador de la encuesta.
 2. *Sesgo de redacción:* La redacción de la pregunta hace que el entrevistado tenga cierta tendencia a escoger una determinada respuesta.
 3. *Sesgos debidos al orden de las respuestas, (Sesgo de primacía, Sesgo de recencia):* Dependiendo del tipo de encuesta, el orden en el que se ofrecen las respuestas a los entrevistados puede hacer que se seleccione preferentemente alguna de ellas. En las encuestas en las que los entrevistados no pueden visualizar las opciones, existe una tendencia a seleccionar la primera alternativa que se mencione (*efecto de primacía*) o la última (*efecto de recencia*), con independencia del contenido de la respuesta en sí.
 4. *Sesgo de contexto:* Los encuestados suelen interpretar los cuestionarios como preguntas enlazadas una tras otra y construyen relaciones internas entre las diferentes preguntas, de tal forma, que las respuestas de una pregunta suelen estar influenciadas en las preguntas anteriores y posteriores, es decir, en el *contexto de la pregunta*.
 5. *Sesgo de deseabilidad social:* En los cuestionarios que se pregunta por aspectos sensibles que pueden inhibir al encuestado a dar una respuesta veraz, en muchos casos, lo que se obtiene no es la opinión real que tiene el entrevistado sino respuestas "políticamente correctas" o socialmente deseables. También aparece este tipo de sesgo cuando elige una respuesta que él cree que es la que se espera que dé, es decir, cuando el encuestado se quiere ajustar a las expectativas o demandas del investigador.
 6. *Sesgo de aquiescencia:* Este sesgo se produce cuando el encuestado elige la respuesta afirmativa o que el entrevistado cree que supone un mayor consenso.
 7. *Sesgo de repetición de respuestas:* Cuando hay muchas preguntas seguidas que tienen las mismas opciones de respuesta, los entrevistados pueden llegar a mostrar una tendencia a repetir las mismas respuestas pregunta tras pregunta.
 8. *Sesgo de aversión de extremos:* En las preguntas con respuestas categorizadas, los encuestados tienden a evitar los extremos, en particular si estos son absolutos (no suelen escoger opciones como "siempre" o "nunca", sino "casi siempre" o "casi nunca").

Estos sesgos han sido ampliamente estudiados [2, 20, 61] y existen muchos trabajos para evitar en la medida de lo posible la influencia de

estas fuentes de error.

Además de los anteriores sesgos, que aparecen debidos al diseño del cuestionario y de las preguntas, existen otras fuentes de error adicionales, introducidas por otras causas diferentes a las anteriormente citadas, entre las que cabe destacar:

ERRORES POBLACIONALES Están debido a una mala selección de la muestra de la población sobre la cual se realizará el estudio. Suelen llamarse *Sesgo por selección* o *Sesgo poblacional*.

ERRORES POR NO RESPUESTA En los casos en los que no se responde el cuestionario completo, o a una pregunta determinada de dicho cuestionario, hay que determinar el porqué. Puede ser por desconocimiento de la temática que se pregunta, por negarse a compartir la información de dicha cuestión, por omisión involuntaria, etc. Todas aquellas preguntas que no son respondidas por los encuestados, pueden alterar el resultado promedio completo, o bien, dar una excesiva importancia a las respuestas proporcionadas por otros encuestados, desplazando el resultado hacia la opinión personal de uno de esos encuestados y no hacia la opinión global, a lo que se le llama *Sesgo por no respuesta*.

ERRORES DE FINALIDAD Cuando el encuestado conoce (o cree conocer) la finalidad u objetivo asociado una pregunta y adapta su respuesta a dicha finalidad u objetivo. Se denomina *Sesgo por finalidad*.

En la sección 4.6.1, se profundizará sobre el *Sesgo por selección*. El segundo tipo (*Sesgo por no respuesta*) se suele dar en las encuestas de tipo personal o en las que hay que dar una opinión de tipo subjetivo; en estas situaciones, es posible que los entrevistados no se atrevan o no sean capaces de responder a algunas preguntas. El *Sesgo de deseabilidad social* aparecería en los casos en que estas preguntas se respondan, pero con respuestas "políticamente correctas", diferentes de la opinión real de los entrevistados. Otra opción posible es que se produzcan respuestas al azar por desconocimiento del tema, aunque no sería demasiado grave, puesto que no generaría un error sistemático, sino aleatorio, que se podría mitigar con una muestra poblacional suficientemente amplia. El tratamiento de la "no respuesta" en cuestionarios tiene una vital importancia, ya que puede alterar completamente los resultados promedio, por un lado, proporcionando excesiva importancia a aquellos cuestionarios que sí han sido respondidos, y por otra parte, porque se ignora el motivo por el cual no se han contestado estos cuestionarios o preguntas. Para evitar, o al menos, limitar estos efectos, se sugiere introducir en los cuestionarios pilotos, previos al diseño del cuestionario final, una opción de "No Sabe" y otra de "No contesta" pudiendo indicar el motivo. Estas opciones de respuestas *abiertas* permiten a los investigadores contar con más elementos de juicio para un mejor diseño de las preguntas. Existen muchos trabajos científicos [2, 4, 20, 61] que se pueden consultar para ampliar información sobre estos tipos de sesgos y conocer, con mayor detalle, los mecanismos habituales que se pueden utilizar para minimizar el efecto de los mismos.

Finalmente, el *Sesgo por finalidad* es una aportación que realiza el autor

de esta Tesis en vista de los resultados obtenidos en el ámbito de las encuestas en las ingenierías, aunque es fácilmente exportable a prácticamente todos los campos del saber. Este tipo de sesgo se describirá con mayor concreción en el Capítulo 8.

4.6 ELECCIÓN DE LA MUESTRA POBLACIONAL

Una de las fases más importantes en el desarrollo de una encuesta es la elección de un grupo de individuos válido y representativo de la población sobre la que se obtendrán las conclusiones. Esta fase puede ser tan simple como un muestreo aleatorio dentro de un grupo poblacional amplio, si se considera que todos los individuos de dicha población proporcionan información con una validez similar, y que no se presentan influencias importantes en sus respuestas. Sin embargo, para encuestas de opinión en las que las respuestas de los individuos pueden estar muy influenciadas por su estatus, condición social, situación geográfica, etc., la elección de los individuos para lograr una muestra poblacional no sesgada puede llegar a significar un escollo considerable.

Hay ciertas cuestiones que hay que determinar a la hora de escoger una muestra poblacional [61]:

- El método de selección de los individuos de la población (tipo de muestreo).
- El tamaño de la muestra poblacional.
- El grado de fiabilidad de las conclusiones en función de la muestra poblacional escogida (probabilidad del error).

4.6.1 Selección de individuos

Una selección incorrecta de la muestra poblacional puede provocar errores posteriores a la hora de estimar las correspondientes opiniones de la población. Por ejemplo, si la muestra perteneciese a un nicho poblacional poco representativo del total de la población, las respuestas y las conclusiones obtenidas no serían válidas, ya que habrían sufrido *Sesgo por selección*. Según Lagares y Puerto [61], de manera generalizada, para evitar este sesgo se realiza una criba previa en la que se indica cuántos individuos se van a tomar, por ejemplo, de cada determinada edad, en cada zona, y con qué características en particular. Para realizar esta criba *a priori* se aplican unas técnicas, denominadas “de muestreo”, que se exponen a continuación:

- A. Muestreo al azar (con y sin reemplazamiento).
- B. Muestreo estratificado.
- C. Muestreo por conglomerados o por racimos.
- D. Muestreo sistemático.

También se pueden dar mezclas entre ellas, como por ejemplo, en el caso del muestreo estratificado por conglomerados, en el que la población se subdivide en distintos conglomerados lo más homogéneos entre sí que a su vez se tratan de manera estratificada (según diferentes conceptos geográficos, sociales, etc.).

Sin embargo, no siempre se puede aplicar una criba previa a la encuesta, ya que a veces se desconoce la clasificación de cada individuo dentro de la muestra poblacional, por lo que para estos casos, lo que se realiza es un descarte *a posteriori* de aquellas respuestas que se desvíen en exceso del comportamiento medio, tratándolas como valores anómalos o extremos.

Tamaño de la muestra

Aunque se sabe que cuanto mayor sea el tamaño de la muestra poblacional, menor será el error inferido, factores como el coste económico, la disponibilidad de tiempo, etc., suelen poner limitaciones a la hora de fijar el valor de este parámetro. Por ello, previamente se debe conocer el tamaño mínimo que debería tener la muestra poblacional para que las conclusiones que se extraigan de las respuestas de los entrevistados sean representativas de la opinión de toda la población. Existen fórmulas que determinan estadísticamente este valor, en función del error mínimo admisible, de si es conocida o no la varianza poblacional, de si se conoce el número de individuos de toda la población, etc. En general, el tamaño de la muestra viene determinado por el nivel de confianza que se pueda obtener de las conclusiones utilizando un determinado tamaño de muestra poblacional. De tal forma que cuanto mayor precisión se requiera para las conclusiones, mayor deberá ser el tamaño de la muestra exigida.

Grado de fiabilidad de los resultados

Como se ha visto en apartados anteriores, todas las decisiones que se toman con respecto al diseño de una encuesta se realizan teniendo en cuenta las conclusiones que se pretendan obtener. En particular, si se desean unas conclusiones muy precisas, será necesaria una muestra amplia, en la que los individuos se hayan seleccionado con una precisión elevada, rechazando aquellos que pudieran falsear los resultados generales. En cambio, si simplemente son admisibles unos resultados que muestren *grosso modo* los aspectos principales de una determinada población, bastará con una muestra de tamaño reducido y un proceso de selección de individuos no excesivamente exigente.

Sin embargo, puede que en algunas ocasiones no exista la posibilidad de elegir a los individuos que formarán parte de la muestra. En estos casos, es interesante calcular el grado de fiabilidad de los resultados obtenidos en función del tamaño de la selección de la muestra poblacional.

Para cada una de las variables cuyo valor se intenta obtener mediante la encuesta, su grado de fiabilidad se puede determinar como un valor único o como un intervalo de confianza. En ambos casos, se realiza una estimación de la distribución estadística que modela cada medida observada, y en función de dicha distribución, se determina el valor (o

el intervalo de valores) que delimita el límite que permitiría identificar un valor concreto de esa medida dentro de un rango válido o aceptable para dicha variable. Las distribuciones estadísticas más utilizadas para este propósito son la distribución Normal o Gaussiana, la χ^2 , la uniforme, la “t” de Student, la “F” de Snedecor, etc. De entre ellas, se escogerá la que mejor se ajuste a los valores de cada variable.

4.7 DISEÑO DE UN CUESTIONARIO

COMO se ha mostrado en la Sección 4.1, el diseño del cuestionario es sólo uno de los pasos del proceso de estudio de investigación mediante cuestionarios. Sin embargo, aunque todos los pasos citados con anterioridad son imprescindibles para obtener datos fiables y útiles, el desarrollo del cuestionario es la etapa del proceso que más esfuerzo requiere y en la que mayor cantidad de decisiones se han de tomar. Así pues, suponiendo que el resto de pasos necesarios para la realización de una encuesta de opinión se han realizado previamente, el diseño de un cuestionario se puede dividir en tres apartados:

1. Determinar qué variable se controla con cada pregunta que se vaya a realizar.
2. Seleccionar la tipología de cada una de las preguntas y especificar su redacción.
3. Diseñar la secuencia de preguntas y la distribución gráfica general del cuestionario.

4.7.1 Dificultades para construir buenas preguntas

Dentro del ámbito de las encuestas de opinión hay una obviedad muy clara: *Para obtener buenas respuestas, hay que hacer buenas preguntas a la gente correcta.* A pesar de lo que se muestra en esta afirmación, es conveniente desmenuzarla para entenderla mejor. Las respuestas “buenas” son aquellas estadísticamente relevantes, de las que se pueden obtener deducciones estadísticas que coincidan con la población completa. En cuanto a la gente “correcta”, se entiende que deben ser individuos de una muestra poblacional representativa.

Por tanto, reescribiendo la frase anterior se puede obtener una afirmación menos “directa” aunque más correcta: *Para obtener respuestas que estadísticamente proporcionen información relevante, hay que saber hacer buenas preguntas a los individuos de una muestra poblacional correctamente escogida y suficientemente representativa.*

Lo único que ha faltado por concretar es qué se entiende por BUENAS PREGUNTAS. La respuesta a esta pregunta es bastante complicada, ya que *a priori* no es posible determinar si una pregunta es buena o es mala. Sin embargo, existen bastantes trabajos [13, 46, 76] en los que se recomiendan normas generales para la redacción de las preguntas, principalmente para encuestas de tipo social. Ejemplos de estas recomendaciones son:

- Realizar preguntas para obtener una de las respuestas objeto de estudio, no caer en la tentación de preguntar “porque es intere-

Las buenas respuestas provienen de buenas preguntas a la gente correcta.

sante saber algo". Este tipo de preguntas suelen confundir a los entrevistados, descentrándolos del cometido principal.

- Limitar la longitud de los cuestionarios, ya que los cuestionarios excesivamente largos no se suelen contestar completamente por parte de los entrevistados, o bien, al alargarse el tiempo, no contienen respuestas muy fiables.
- Comenzar con preguntas sencillas y atrayentes para los entrevistados, a las que no tengan que dedicar excesivo tiempo para pensar la respuesta. Esta estrategia les anima a continuar. Si las preguntas iniciales son demasiado críticas, o implican que los entrevistados les dediquen mucho tiempo para pensar las respuestas, es muy probable que renuncien a continuar con el cuestionario.
- Utilizar preguntas directas y con un lenguaje simple. Las preguntas enrevesadas, con un lenguaje excesivamente técnico, o por el contrario, utilizando jerga, habitualmente no son comprendidas por los entrevistados, y por lo tanto, sus respuestas suelen ser poco fiables.

Las recomendaciones anteriores son de tipo general. Sin embargo, observando determinadas preguntas se pueden deducir otras recomendaciones más específicas para la redacción de las preguntas en particular. A continuación, se propone un conjunto de preguntas *mal planteadas*, junto con la explicación de porqué se considera que dichas preguntas no están bien formuladas y cómo rehacerlas para obtener BUENAS PREGUNTAS.

EVITAR LA AMBIGÜEDAD: Estos ejemplos muestran que no se deben diseñar preguntas que puedan tener múltiples interpretaciones o preguntas que fuercen al encuestado a responder de una determinada manera específica.

EJEMPLO:

¿Qué te gusta más, el sabor cítrico o la textura crujiente del helado?

COMENTARIO:

Se dan demasiadas combinaciones para las respuestas: que al entrevistado le guste el sabor cítrico pero no la textura crujiente, que le gusten ambos, que no le gusten ninguna de las dos, que le guste la textura crujiente pero no el sabor, o incluso que le gusten ambos, pero por separado.

Preguntar por un único aspecto en cada pregunta.

EJEMPLO:

¿Qué sistema operativo utilizas en tu PC?

Windows XP o Windows Vista.

COMENTARIO:

No es una buena pregunta, puesto que se pueden dar otros muchos tipos de sistemas operativos: ninguno, Linux, Solaris, MacOS, otras versiones diferentes de MS Windows, etc.

Cubrir todas las respuestas posibles para cada pregunta.

EJEMPLO:

¿Qué lugares te gustaría visitar como turista?

Paris, Madrid o Francia.

COMENTARIO:

Las respuestas no han sido correctamente seleccionadas y pueden presentar dudas a los entrevistados, ya que el que se escoja visitar Francia no es excluyente para escoger Paris.

Selecciona respuestas mutuamente excluyentes entre sí.

Con recomendaciones de este tipo se pretende prevenir la ambigüedad en las preguntas. Evitando este problema, se trata de obtener de los entrevistados respuestas veraces. La veracidad de las respuestas tiene que venir de dos frentes: los entrevistados deben entender perfectamente lo que se les pregunta y han de tener opinión propia sobre la cuestión.

PREGUNTAS CAPCIOSAS Y MAL PLANTEADAS: Este contexto abarca tanto las **PREGUNTAS CAPCIOSAS**, que están diseñadas con una intencionalidad predeterminada por parte del encuestador, como las preguntas que, por su redacción, provocan una influencia no deseada en la respuesta de los entrevistados, aunque, *a priori* no han sido diseñadas con ninguna intencionalidad. Estas recomendaciones pretenden la independencia de las respuestas de los encuestados. Ésta proviene principalmente de la necesidad de separar el pensamiento del entrevistado de las ideas preconcebidas del entrevistador: el entrevistado debe responder por sí mismo, sin ayuda, guía o consejo del entrevistador. Es evidente que para que los resultados de la encuesta sean válidos, el entrevistador debe evitar expresar cualquier opinión personal o cualquier resultado general, ya que estos comentarios, sugerencias o datos adicionales suelen condicionar las respuestas. El primer ejemplo introduce información propia del entrevistador en la pregunta:

EJEMPLO:

En lo concerniente a Fórmula 1, sabiendo que Ferrari es la mejor escudería, ¿cree Vd. que este año conseguirán revalidar el título?

COMENTARIO:

Esta se puede considerar una **PREGUNTA CAPCIOSA**: la información de que Ferrari es la mejor escudería es propia del entrevistador y no proporciona información útil para el entrevistado, por lo que influye (o al menos podría influir) en la respuesta del entrevistado. Una pregunta más aséptica y neutral sería: *¿Cree Vd. que Ferrari conseguirá revalidar el título de Fórmula 1 este año?*

No introducir opiniones personales del entrevistador en el cuerpo de las preguntas.

En este ejemplo, se añade información general que puede condicionar la respuesta:

EJEMPLO:

¿Está Vd. de acuerdo con la mayoría de la gente que opina que la economía va bien?

COMENTARIO:

Esta pregunta, *a priori*, no puede ser considerada una PREGUNTA CAPCIOSA, sin embargo, está mal diseñada, ya que podría inducir un sesgo de *deseabilidad social*. Al introducir información sobre qué ha contestado “la mayoría de la gente”, se está proporcionando al entrevistado datos sobre las respuestas que ya se han obtenido. Esto, en cierto modo, obliga al entrevistado a dar una respuesta similar a los demás. La pregunta estaría mejor formulada como *¿Cree Vd. que la economía va bien?*

No añadir datos de opiniones de tipo general que puedan influir en la respuesta.

PREGUNTAS CON EXCESO DE INFORMACIÓN: Existe otro tipo de condicionamiento de las respuestas, que es mucho más sutil, y que en la mayoría de los casos no se realiza conscientemente por parte del entrevistador. Suele aparecer cuando se incluye información adicional, datos o descripciones para explicar el entorno en el que se ubica una determinada pregunta. Esta información puede llegar a modificar o alterar la independencia de los entrevistados, y resulta muy difícil, a veces imposible, evitar este efecto.

EJEMPLO:

Las tres grandes compañías de desarrollo de videoconsolas, Sony, Microsoft y Nintendo, han sacado al mercado diferentes consolas de última generación, cada una de éstas con distintos precios (de manera aproximada, PS3 de Sony: 400 €; Xbox360 de Microsoft, 300 €; Wii de Nintendo, 250 €). ¿Cuál considera que es la mejor de estas tres consolas en relación calidad-precio?

COMENTARIO:

El entrevistador ha proporcionado los precios aproximados de las tres videoconsolas como datos para que el entrevistado posea más información y pueda opinar con mayor fiabilidad. Al asegurarse el entrevistador de que el entrevistado posee información sobre el precio, éste presupone que el entrevistado podrá opinar mejor. Sin embargo, al dar información sobre el precio puede que no se obtenga la opinión directa del entrevistado, ya que dicha información podría modificar la opinión prefijada que el entrevistado pudiese tener, lo cual puede ser aceptable o no según los casos. Por ejemplo, si no se hubiesen incluido los precios y el entrevistado no tuviese una información exacta del precio de cada consola, o si su información fuese errónea, probablemente su respuesta sería diferente.

Sin ser erróneo, añadir ciertos datos, que el entrevistador debe presuponer que conocen los entrevistados, puede influir en las respuestas.

4.7.2 Buenas preguntas con buenas respuestas

De manera análoga al planteamiento de las preguntas, si se ofrecen respuestas acotadas, su redacción debe seguir una serie de recomendaciones para que no influyan en la decisión de los entrevistados. En

algunos casos, la inclusión de información adicional en las respuestas, ya sea explícita o subjetiva, puede influir de forma capciosa y malintencionada en los entrevistados, de manera análoga a la que se ha comentado anteriormente para las preguntas. En otros casos, puede que sea la información adicional incluida en las preguntas la que condicione la selección de una de las posibles respuestas por parte del entrevistado.

INFORMACIÓN EXPLÍCITA EN LAS RESPUESTAS: La inclusión de información detallada en las respuestas puede influir en la selección que haga el encuestado.

EJEMPLO:

El gobierno pretende aumentar el impuesto de circulación con el objetivo de concienciar a la población acerca del cambio climático. Según Ud. ¿qué vehículos deberían pagar mayores impuestos por dicho concepto? (Puede seleccionar más de una).

- a. Los vehículos que más contaminan (4 × 4 y todoterrenos).*
- b. Todos los vehículos industriales (camiones, furgonetas y similares).*
- c. Todos los vehículos por igual.*
- d. No deberían aumentar el impuesto de circulación por ese motivo.*

COMENTARIO:

La primera respuesta añade explícitamente ejemplos de cuáles considera el entrevistador que son los vehículos que más contaminan (4 × 4 y todoterrenos). Esta opinión del entrevistador no debería estar incluida puesto que puede estar limitando la elección del entrevistado.

INFORMACIÓN SUBJETIVA EN LAS RESPUESTAS: También se debe evitar el uso de información adicional en las respuestas que pueda ser interpretada de manera subjetiva por los entrevistados. Ocurre especialmente al utilizar imágenes, videos, sonidos, etc., aunque también puede producirse en cualquier tipo de respuesta, si se expresa mediante términos vagos que el usuario pueda interpretar subjetivamente.

No añadir información adicional en las respuestas.

EJEMPLO:

El gobierno está pensando aumentar los fondos de las campañas sociales de apoyo a la familia. ¿Quién cree Vd. que debería sufragar principalmente dichos fondos? (Puede seleccionar más de una).

- a. Los ricos.*
- b. Todas las empresas que tengan más de 1000 empleados.*
- c. Toda la sociedad a través del impuesto de IRPF.*

COMENTARIO:

Aunque la pregunta está bien planteada, entre las respuestas que se ofrecen se incluye una que tiene connotaciones negativas: “los ricos”. Al no determinar de manera explícita qué requisitos debe cumplir una persona para ser considerada rica, esta respuesta es un “cajón de sastre” en el que cada entrevistado puede incluir de manera subjetiva a cualquier persona. Además, el sustantivo “rico” podría suscitar en los entrevistados ideas peyorativas con respecto a la relación del compromiso social global frente a la situación económica de dichas personas.

No utilizar expresiones que tengan connotaciones subjetivas adicionales, ya sean positivas o negativas.

SUBJETIVIDAD COMPARADA: Aparece en preguntas en las que se comparan dos o más posibilidades entre sí. En estas preguntas no se solicita la opinión de manera absoluta sobre un aspecto en particular, sino en relación (comparándolo) con otro.

EJEMPLO:

Con respecto a la calidad de la comida en los locales de comida rápida, seleccione con cuál de las siguientes afirmaciones está Vd. de acuerdo:

- a. Es peor que la de los locales de comida tradicional.*
- b. Tienen la misma calidad que la de los locales de comida tradicional.*
- c. Es mejor que la de los locales de comida tradicional.*

COMENTARIO:

Es probable que el entrevistador pretenda, con esta comparación, averiguar la opinión de los entrevistados acerca de la calidad de la comida en los locales de comida rápida. Sin embargo, dicha comparación no indica lo buena o mala que el entrevistado considera en sí la calidad de la comida de los locales de comida rápida, sino si la considera mejor o peor que los de comida tradicional. El entrevistado puede considerar que los dos tienen muy mala calidad, aunque uno de los dos sea aún peor que el otro; o justo al revés, que los dos tengan buena calidad, aunque uno de ellos tenga mejor calidad que el otro. Es decir, que la subjetividad comparativa que se introduce no se resuelve con una sola pregunta. Habrá que hacer más preguntas para catalogar la calidad individual de cada uno de los dos.

Las respuestas subjetivas comparativas no garantizan conclusiones con catalogación absoluta, sino relativa.

4.8 ANÁLISIS DEL CUESTIONARIO

Hasta ahora se ha expuesto una serie de recomendaciones para diseñar un buen cuestionario desde el punto de vista lingüístico y sociológico. Sin embargo, para determinar la corrección del cuestionario hay que realizar un análisis estadístico del mismo.

Desde dicho ámbito, un cuestionario pretende realizar una medida indirecta de una o varias variables estadísticas no mesurables directamente. Debido a esta capacidad de medida, a los cuestionarios también se les suele denominar **INSTRUMENTOS**. Las preguntas individuales utilizadas dentro de un cuestionario se llaman **ITEMS**, que se agrupan en **DOMINIOS** (también referidos en la literatura científica como **SUBESCALAS**). Habitualmente cada **DOMINIO** tiene asociada una variable estadística diferente.

Los diseñadores suelen agrupar los **ITEMS** en **DOMINIOS** de tres maneras diferentes:

- A. Basándose exclusivamente en el juicio de uno o varios expertos.
- B. Basándose exclusivamente en la exploración preliminar de los Componentes Principales de cada pregunta del cuestionario (mediante **PCA**).
- C. Basándose en el juicio del experto, refinado o reforzado mediante una exploración confirmatoria con **PCA**.

El análisis estadístico **PCA** busca la correlación de las respuestas entre los diferentes **ITEMS** del cuestionario y trata de comprobar si las respuestas sostienen estadísticamente la distribución de **DOMINIOS** que se ha realizado. Si un conjunto de **ITEMS** presentan una alta correlación, entonces probablemente midan la misma variable estadística y por tanto, es posible asignarlos a un mismo **DOMINIO**. Por el contrario, si los **ITEMS** muestran una baja correlación, entonces es probable que midan variables estadísticas diferentes y que no pertenezcan al mismo **DOMINIO**. Finalmente, existen **ITEMS** que podrían correlacionarse con varias variables estadísticas diferentes, por lo que podrían incluirse en diferentes **DOMINIOS**. Por tanto, los componentes principales que se obtienen tras la aplicación de **PCA** al cuestionario representan estimaciones de aquellas variables que estadísticamente son independientes.

4.8.1 Validación del cuestionario

Uno de los aspectos más deseados, y a la vez más complejos, cuando se trabaja con cuestionarios es la garantía de su **VALIDEZ**. La **VALIDEZ** de un cuestionario se puede observar desde tres evidencias diferentes:

- A. Validez de **CONSTRUCTO**.
- B. Validez de contenido.
- C. Validez de criterio.

La *validez de* **CONSTRUCTO** se utiliza para demostrar que las respuestas obtenidas a través del cuestionario se ajustan al mismo número de **DOMINIOS** incluidos en el diseño del cuestionario. Para comprobarla se

INSTRUMENTOS:

Cuestionarios

ITEMS: Preguntas individuales de un cuestionario

DOMINIOS:

Agrupaciones de preguntas de un cuestionario que son relativas a un mismo aspecto

PCA: En inglés,

PRINCIPAL

COMPONENT

ANALYSIS. *Análisis de Componentes*

Principales

VALIDEZ: *Determina si un cuestionario realmente mide lo que afirma medir.*

CONSTRUCTO: *Del latín, construcción teórica de una variable compleja, compuesta por varios dominios*

aplican tanto PCA (validez factorial) como intercorrelaciones (validaciones de convergencia y de divergencia) sobre el total de las respuestas obtenidas.

Mediante la *validez de contenido* se justifica la elección de los diferentes ITEMS utilizados dentro del cuestionario, así como la de las escalas de sus respuestas. Esta validez se comprueba mediante el consenso de los expertos, que deben estar de acuerdo en que tanto las preguntas como sus posibles respuestas son acordes al ámbito del cuestionario.

Finalmente, la *validez de criterio* comprueba si el INSTRUMENTO es suficientemente sensible y específico como para ser útil en la medida del concepto para el que se construyó el cuestionario. Para ello, se compara la capacidad de discriminación del cuestionario frente a un **GOLD STANDARD**. A veces, también se pueden utilizar medidas indirectas. Por ejemplo, para determinar si un paciente sufre BAJA VISIÓN se pueden utilizar instrumentos que midan su agudeza visual o su campo de visión. Es evidente que si al realizar una medición del campo de visión de un paciente, éste es menor del 20%, el paciente sufre BAJA VISIÓN. Por otra parte, también se podría llegar a dicha conclusión utilizando un cuestionario que vaya realizando preguntas hasta determinar de forma indirecta si un paciente sufre BAJA VISIÓN. En este caso, se podría asegurar la *validez de criterio* de dicho cuestionario si sus conclusiones acerca de la BAJA VISIÓN de todos los pacientes encuestados coinciden con los resultados proporcionados por los instrumentos de medida del campo visual.

GOLD STANDARD:
Una medida objetiva e inequívoca que permita determinar de manera perfecta y sin error la variable que se desea observar en cada individuo. Muchas veces, no existe o es muy compleja de conseguir.

4.8.2 *Fiabilidad del cuestionario*

Otro aspecto de gran importancia que hay que tener en cuenta a la hora de aplicar un cuestionario es la fiabilidad de las medidas que proporciona. La **FIABILIDAD** se puede comprobar desde varios puntos de vista:

FIABILIDAD:
Capacidad de un cuestionario de obtener resultados análogos con otras muestras poblacionales o en otros instantes temporales.

FIABILIDAD INTERNA Comprueba si los ITEMS muestran consistencia interna, es decir, si todos los ITEMS que están agrupados en un mismo DOMINIO presentan una alta correlación.

REAPLICABILIDAD Comprueba si los resultados obtenidos al volver a aplicar los cuestionarios son estadísticamente similares a los obtenidos al aplicarlos por primera vez (también llamado efecto "test-retest").

En ambos casos se utilizan métodos como el cálculo del estadístico α -Cronbach, medidas de la correlación intraclases, o estadísticos equivalentes como el *índice de concordancia de Kendall*, el coeficiente de concordancia de las correlaciones o el coeficiente *Kappa*.

4.8.3 *Alfa de Cronbach*

El estadístico α -Cronbach se utiliza como una de las medidas de fiabilidad o consistencia interna de un cuestionario psicométrico sobre

una muestra específica de encuestados. La fórmula asociada a dicho estadístico se muestra en (4.1).

$$\alpha_{\text{cronbach}} = \frac{K}{K-1} \cdot \left(1 - \frac{\sum_{i=0}^n \sigma_{Y_i}^2}{\sigma_X^2} \right) \quad (4.1)$$

Donde K representa el número de preguntas o ítems que componen el cuestionario, $\sigma_{Y_i}^2$ es la varianza de las respuestas de los encuestados teniendo en cuenta únicamente la pregunta i -ésima y con σ_X^2 se indica la varianza de la totalidad de las preguntas para el mismo conjunto de encuestados indicado con anterioridad.

Este estadístico se puede utilizar para calcular la fiabilidad, en términos de consistencia interna, de una encuesta siempre y cuando ésta presente las siguientes características:

1. La encuesta está compuesta por preguntas que se combinan entre sí mediante sumas para proporcionar una puntuación global.
2. Todas las preguntas miden la misma característica con la misma proporcionalidad. Es decir, no existen preguntas en la que un valor menor signifique que la característica medida se presenta en mayor grado. Todas las preguntas deben mostrar una correlación lineal positiva en la graduación de las respuestas.
3. Todas las preguntas deben estar en la misma escala de graduación. No es admisible tener preguntas valuadas en un rango continuo $[0, 10]$ y otras en un rango dicotómico $0 - 1$.

En caso de que las encuestas no cumplan todos estos requisitos, habrá que hacer una recodificación de los datos para adaptarlos a estas exigencias.

Descripción cualitativa del valor numérico del estadístico

El estadístico α -Cronbach presenta una doble interpretación, desde un punto de vista sociológico y desde un punto de vista matemático. En el plano sociológico, lo deseable para crear una encuesta fiable es que todas las preguntas que componen el cuestionario estén muy correlacionadas entre sí. Este estadístico permite una medida cuantitativa del grado de cohesión que muestran todas las preguntas entre sí. Desde un punto de matemático, el nivel máximo de correlación se alcanza cuando las respuestas de las preguntas Y_1, Y_2, \dots, Y_K son idénticas. En ese caso, se cumple que $\sigma_{Y_j}^2 = K^2 \cdot \sigma_X^2$, y también que, $\sum_{i=1}^K \sigma_i^2 = K \cdot \sigma_X^2$, por lo que (4.1) se puede simplificar, obteniéndose como resultado en dicho caso el valor 1. El opuesto sucede cuando cada pregunta es completamente independiente de las demás y, por tanto, no preguntan sobre un ámbito común. Para ese supuesto, al ser cada pregunta independiente, la varianza total es igual a la suma de la varianza de cada pregunta, esto es: $\sigma_X^2 = \sum_{i=1}^K \sigma_{Y_i}^2$. Esta formulación hace que (4.1) se iguale a 0. Finalmente, en los casos en los que existan preguntas que estén inversamente correlacionados, α -Cronbach puede llegar a alcanzar valores negativos.

Este estadístico muestra un comportamiento matemático en el que se

Un α -Cronbach cercano a 1 muestra una alta fiabilidad interna.

sobrevalora mucho el número de preguntas. De tal manera que para un cuestionario con un número K de preguntas se obtiene un cierto valor α -Cronbach, si se añaden preguntas a dicho cuestionario, independientemente del grado de relación de estas nuevas preguntas con las anteriores, el valor α -Cronbach aumenta. Por tanto, cuestionarios que presentan valores muy elevados del estadístico α -Cronbach han de tomarse en cuenta con cierta prudencia, puesto que esos valores podrían provenir de una sobrevaloración por un número elevado de preguntas más que por presentar una gran consistencia entre las preguntas. Para dar validez a estos resultados hay que analizar otros estadísticos, que reafirmen la consistencia de los resultados.

4.8.4 Coeficiente de Correlación Intraclase

El ICC [99] es un estadístico que mide la coherencia de las respuestas con diversos grados de libertad. Ha sido utilizado habitualmente para mostrar el grado de correlación en las pruebas *test-retest* y en las pruebas de concordancia entre múltiples mediciones de una misma variable. También se ha utilizado para estimar la fiabilidad entre las medidas de distintos evaluadores sobre conjuntos de sujetos u observadores diferentes, o bien, sobre el mismo conjunto de sujetos fijo pero con diferentes instrumentos de medida. La interpretación matemática del ICC se deriva de un modelo de análisis de varianza con efectos mixtos. La varianza total entre las mediciones se debe a tres factores: diferencias de interpretación entre observadores y evaluadores de dichos observadores (o bien entre distintos instrumentos de medida), las diferencias entre las diferentes variables medidas y los residuos de error en la varianza que no se explican por los otros casos. En general, ICC se calcula como la relación de las varianzas entre las respuestas de los sujetos y las de los residuos de error:

ICC: En inglés, INTRACLASS CORRELATION COEFFICIENT, Coeficiente de Correlación Intraclase

$$ICC = \frac{\sigma_{\text{sujetos}}^2}{\sigma_{\text{sujetos}}^2 + \sigma_{\text{error}}^2} \quad (4.2)$$

Shrout y Fleiss definieron [99] tres formulaciones de ICC, en las que las varianzas se estiman mediante análisis ANOVA.

ICC(1, k) Conjunto de evaluadores seleccionado aleatoriamente. Conjunto de sujetos seleccionados aleatoriamente. Selección aleatoria de un evaluador para cada sujeto. Diseño anidado.

Se debe aplicar sobre cuestionarios con diseño anidado de diferentes sujetos bajo cada evaluador, en el que el conjunto de evaluadores (o el conjunto de instrumentos de medida) se seleccionan aleatoriamente y el conjunto de sujetos u observadores se selecciona también de manera aleatoria. La fórmula para calcular ICC(1, k) es:

$$ICC(1, k) = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma_i^2}{k}} \quad (4.3)$$

Ésta calcula el grado de acuerdo en la respuesta de la media de k evaluadores, donde σ_p^2 representa la varianza de la población y

σ_i^2 indica la varianza de las respuestas intra-personas, estimadas mediante análisis ANOVA.

ICC(2,k) Conjunto de evaluadores seleccionado aleatoriamente. Conjunto de sujetos seleccionados aleatoriamente. Diseño cruzado. Se utiliza sobre cuestionarios con diseño cruzado de sujetos y evaluadores, en los que tanto el conjunto de evaluadores como el conjunto de sujetos se seleccionan de manera aleatoria. La fórmula para calcular ICC(2,k) es:

$$\text{ICC}(2, k) = \frac{\sigma_p^2}{\sigma_p^2 + \frac{(\sigma_r^2 + \sigma_e^2)}{k}} \quad (4.4)$$

Esta formulación calcula el grado de acuerdo medio a lo largo de un conjunto aleatorio de k evaluadores, donde σ_p^2 representa la varianza de la población, σ_r^2 representa la varianza sobre el conjunto de evaluadores y σ_e^2 indica la varianza de los residuos de error.

ICC(3,k) Conjunto de evaluadores fijo. Conjunto de sujetos seleccionados aleatoriamente. Diseño cruzado. Aplicado con cuestionarios con diseño cruzado de sujetos y evaluadores. Para el cálculo de este estadístico, tanto el conjunto de evaluadores (o el conjunto de instrumentos de medida) como el conjunto de observadores, se seleccionan de forma aleatoria. La fórmula para calcular ICC(3,k) es:

$$\text{ICC}(3, k) = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma_e^2}{k}} \quad (4.5)$$

Gracias a la cual se calcula la consistencia/fiabilidad de la respuesta media a lo largo de un conjunto fijo de k evaluadores.

Estas tres formulaciones producen resultados totalmente diferentes y no son intercambiables, por lo tanto, la elección correcta del estadístico es fundamental. Una descripción extensa y pormenorizada de estos ha sido realizada por McGraw y Wong [70].

Descripción cualitativa del valor numérico del estadístico

El valor del estadístico ICC se suele analizar utilizando la Tab. 4.1 para determinar, de manera cualitativa, el grado de acuerdo entre los diversos evaluadores y/o sujetos, o bien el grado de fiabilidad de las respuestas medias de los evaluadores y/o sujetos.

ICC	VALORACIÓN
< 0.1	Nula
0.11 – 0.3	Pobre
0.31 – 0.5	Mediocre
0.51 – 0.7	Moderado
0.71 – 0.9	Bueno
> 0.91	Excelente

Tabla 4.1
Valoración cualitativa
del ICC.

4.9 CONCLUSIONES

Se ha utilizado este capítulo para presentar los estudios de investigación mediante cuestionarios, mostrando las fases que hay que seguir para desarrollarlo. También se han indicado los diversos métodos para la recolección de las respuestas de los entrevistados, las limitaciones que presentan tanto los cuestionarios, en general, como estos mecanismos de recolección de respuestas, en particular.

Se ha dedicado una sección completa a la descripción y estudio de las fuentes de errores en los cuestionarios. En este contexto, se ha identificado un tipo de error, el *sesgo por finalidad*, del que no se tiene constancia que hubiera sido estudiado o descrito en la literatura científica con anterioridad a esta Tesis. Debido al interés que esta nueva aportación pudiera suscitar, se dedicará un apartado en el Cap. 8 a describir y estudiar este *sesgo*, y a proponer posibles soluciones al mismo para eliminar o limitar su efecto en lo posible.

También es importante hacer notar el trabajo desarrollado en la Sec. 4.7, en la que no sólo se han indicado las recomendaciones propuestas por otros autores para el diseño de las preguntas de los cuestionarios, sino que se han propuesto ejemplos que muestran preguntas erróneas, con la explicación detallada de porqué se pueden considerar que son erróneas y posibles soluciones, en cada caso.

El Capítulo 2 sirvió para exponer los fundamentos de diferentes metodologías en el ámbito del procesamiento de imágenes y de algunos algoritmos de extracción de bordes; en el Capítulo 3 se mostró que para determinar la calidad de los resultados de dichos algoritmos, los resultados más precisos se obtienen utilizando la evaluación subjetiva. En este Capítulo 4, se ha profundizado en los cuestionarios y en cómo se deben diseñar para que estos sean capaces de obtener medidas correctas de variables subjetivas dentro de una población de individuos, como la calidad percibida por los usuarios. El próximo capítulo se realizará una exposición de diferentes estudios de investigación basados en cuestionarios para pacientes con BAJA VISIÓN, que permitirá determinar los DOMINIOS más relevantes para estos pacientes.

CUESTIONARIOS SOBRE BAJA VISIÓN

ÍNDICE

5.1	Investigación sobre la BAJA VISIÓN mediante cuestionarios	90
5.1.1	Visual Function Index	90
5.1.2	Rand Questionnaire to Assess Functional Problems of the Visually Impaired	91
5.1.3	Visual Function after Pan-Retinal Photocoagulation	91
5.1.4	Visual Status Inventory	92
5.1.5	Lowe's Visual Function	92
5.1.6	Questionnaire for Functional Assessment of Low Vision	92
5.1.7	Visual Activities Questionnaire	93
5.1.8	Visual Performance Questionnaire	93
5.1.9	Activities of Daily Vision Scale	94
5.1.10	Assessment of Visual Function-Related Quality of Life	94
5.1.11	14-Item Visual Functioning Index	95
5.1.12	Vision-Related Quality of Life Questionnaire	95
5.1.13	National Eye Institute's Visual Functioning Questionnaire	96
5.1.14	Cuestionario de Calidad Visual del Consejo Argentino de Oftalmología	97
5.2	Dominios relevantes para BAJA VISIÓN y procesamiento de imágenes	98
5.3	Conclusiones	100
5.3.1	Conclusiones de la Parte I	100

Puesto que uno de los ámbitos de actuación de esta Tesis es el de los pacientes con BAJA VISIÓN, a continuación se expone un estudio en el que se muestra qué aspectos se contemplan en los cuestionarios de evaluación de los pacientes de BAJA VISIÓN. Este capítulo se basa en la revisión científica de diversos cuestionarios de evaluación de la función visual de pacientes con BAJA VISIÓN realizada por *Massof et al.* [68]. La mayoría de estos cuestionarios se diseñaron para detectar la BAJA VISIÓN en pacientes, y ninguno trata de medir el grado de acomodación de estos pacientes a imágenes procesadas. Todos los cuestionarios estudiados en dicho artículo son de tipo *lingüístico*, mientras que los que se necesitan para la evaluación de los algoritmos de procesamiento de contornos, que se utilizarán a lo largo de esta Tesis, han de ser de tipo visual. Sin embargo, aunque la revisión de estos cuestionarios no garantizará la VALIDEZ del cuestionario de evaluación de calidad de bordes, puede servir como un buen punto de inicio para el desarrollo de dicho cuestionario, sobre todo para escoger los grupos de imágenes que se utilizarán para su procesamiento.

5.1 INVESTIGACIÓN SOBRE LA BAJA VISIÓN MEDIANTE CUESTIONARIOS

Existen algunos estudios que evalúan tanto la función visual de los pacientes que sufren BAJA VISIÓN como la influencia de su patología en otros aspectos de su vida habitual. Massof *et al.* [68] realizaron un artículo en el que se realizó una revisión científica muy exhaustiva de muchos y diversos cuestionarios de evaluación de la función visual de pacientes con BAJA VISIÓN. Tomando como base dicho trabajo, a continuación, se realizará una exposición de los principales cuestionarios para pacientes con BAJA VISIÓN, haciendo especial hincapié en la descripción de los DOMINIOS de cada cuestionario. Se ha optado por indicar el nombre de cada cuestionario en el idioma original sin traducirlo al español, ya que muchos de estos cuestionarios están ampliamente difundidos con los nombres originales y la traducción de los mismos podría llevar a problemas de identificación y dificultar su localización en otros textos científicos.

5.1.1 Visual Function Index

VFI: VISUAL
FUNCTION INDEX,
Índice de Función
Visual.

Desarrollado por Bernth-Petersen [9], el cuestionario VFI está dividido en 9 ITEMS, que cada uno podría identificarse, a su vez, con un DOMINIO. Cada pregunta con una graduación de posibles respuestas diferentes:

1. Lectura (*Nada, Sólo letras grandes, Letras pequeñas*).
2. Visión a media/larga distancia (*Buena, Moderada, Pobre*).
3. Ver la Televisión (*Sí, No*).
4. Conducción de automóvil/bicicleta (*Sí, No*).
5. Orientación en interiores (*Sí, No*).
6. Orientación en exteriores (*Sí, No*).
7. Actividad "principal": trabajo, tareas del hogar (*Sí, No*).
8. Otras actividades: ocio, ir de compras, etc. (*Sí, No*).
9. Actividades de "cuidado personal": comer, aseo, bañarse, etc. (*Sí, No*).

5.1.2 *Rand Questionnaire to Assess Functional Problems of the Visually Impaired*

Bikson y Bikson [10], de la empresa **Rand Corp.**, diseñaron el cuestionario **RAND-FPVI**, compuesto por 30 ITEMS agrupados en 8 DIMENSIONES:

1. Habilidades para la vida independiente.
2. Orientación general.
3. Movilidad general.
4. Viajar en autobús.
5. Problemas de iluminación.
6. Tareas del hogar.
7. Percepción social.
8. Actividades recreativas.

RAND-FPVI: RAND QUESTIONNAIRE TO ASSESS FUNCTIONAL PROBLEMS OF THE VISUALLY IMPAIRED. Cuestionario Rand para determinar los Problemas Funcionales de los Deficientes Visuales.

En todos los casos, las preguntas tienen 3 posibles respuestas: *Sin afectación*, *Afectado por sus problemas de visión*, *Afectado por otros problemas distintos a los de su visión*.

5.1.3 *Visual Function after Pan-Retinal Photocoagulation*

El cuestionario **VF-PRP** fue desarrollado inicialmente por **Russell et al.** [95] y posteriormente fue refinado por **Kosnik et al.** [60]. Este cuestionario consta de 27 ITEMS que se dividen en 8 DOMINIOS:

1. Velocidad de procesamiento visual.
2. Sensibilidad a la luz.
3. Visión a distancia.
4. Visión cercana.
5. Visión binocular.
6. Visión nocturna.
7. Falta de fluidez visual.
8. Adaptación a la luz.

VF-PRP: VISUAL FUNCTION AFTER PAN-RETINAL PHOTOCOAGULATION. Función Visual tras Fotocoagulación Pan-Retinal.

En todos los casos, los ITEMS presentan a los pacientes respuestas en las que se mide cuándo han tenido dificultad al realizar una determinada tarea en relación con la fotocoagulación pandiabética: *Nunca*, *Antes*, *Justo después*, *Ahora*. En caso de responder *Ahora*, se añade una segunda escala en la que se cataloga el grado de la dificultad: *Suave*, *Moderada*, *Severa*, *Intolerable*.

5.1.4 *Visual Status Inventory*

VSI: VISUAL STATUS INVENTORY, *Inventario de Estado Visual.*

Coren y Hakstian diseñaron y validaron el cuestionario VSI [25] con 52 ITEMS para la estimación de la discapacidad visual en estudios epidemiológicos. Estos ITEMS se distribuyen en tres DOMINIOS:

1. Agudeza visual.
2. Visión en color.
3. Función binocular general.

Las respuestas pretenden medir la frecuencia con la que los pacientes son capaces de realizar una actividad. Éstas están estructuradas utilizando una escala de Lickert de 5 puntos, que cubre desde *Nunca* hasta *Siempre*.

5.1.5 *Lowe's Visual Function*

LVF: LOWE'S VISUAL FUNCTION, *Función Visual de Lowe.*

Lowe desarrolló un cuestionario, denominado genéricamente LVF, que posteriormente fue refinado por Elliott *et al.* [32], para la medición de la función visual para pacientes con cataratas en, al menos, un ojo.

Este cuestionario está compuesto por 14 ITEMS agrupados en 4 DOMINIOS:

1. Movilidad: Caminar por el exterior, cruzar la calle, conducir, etc.
2. Visión cercana: Lectura de libros, lectura de periódicos, etc.
3. Discriminación: Reconocimiento de amigos, lectura del número de autobús, ver la televisión, etc.
4. Calidad de visión: Visión con el ojo derecho, visión con el ojo izquierdo, visión binocular.

Las respuestas se presentan como una escala continua de 10 cm. en la que los pacientes deben situar una marca en la que se identifica el grado de discapacidad que sufren ante cada ITEM.

5.1.6 *Questionnaire for Functional Assessment of Low Vision*

FALV: QUESTIONNAIRE FOR FUNCTIONAL ASSESSMENT OF LOW VISION, *Cuestionario para Determinación Funcional de la BAJA VISIÓN.*

Szlyk *et al.* [107] diseñaron un cuestionario para catalogar de manera funcional el grado de BAJA VISIÓN de los pacientes. Este cuestionario, llamado FALV, está basado en el trabajo de Coren y Hakstian [25] (diseñadores del cuestionario VSI), y en el de Kosnik *et al.* [60] (refinaron el cuestionario VF-PRP). Este cuestionario se ha implementado teniendo en cuenta DOMINIOS funcionales que pudieran ser evaluados también por los instructores de rehabilitación de los pacientes, simplemente observando el comportamiento de los mismos.

Este cuestionario está compuesto por 57 ITEMS agrupados en 4 DOMINIOS:

1. Búsqueda (*Cerca, Cerca por debajo, Cerca por encima, Lejos*).
2. Detección (*Cerca, Cerca por debajo, Cerca por encima, Cerca a la izquierda, Cerca a la derecha, Lejos*).

3. Análisis (*Sí, No*).
4. Seguimiento (*Sí, No*).

5.1.7 *Visual Activities Questionnaire*

Los pacientes de la 3ª edad suelen presentar una disminución de la calidad visual, y Sloane *et al.* [102] quisieron medir dicho nivel de calidad visual de manera indirecta, a través de la capacidad de los pacientes para realizar actividades cotidianas. Esta medición se realizó mediante un cuestionario denominado VAQ.

Este cuestionario está compuesto por 33 ITEMS agrupados en 8 DOMINIOS:

1. Visión periférica.
2. Agudeza visual.
3. Búsqueda visual.
4. Profundidad.
5. Color.
6. Adaptación.
7. Deslumbramientos.
8. Velocidad de procesamiento visual.

VAQ: VISUAL
ACTIVITIES
QUESTIONNAIRE,
Cuestionario de
Actividades Visuales.

Las respuestas a los ITEMS representan la frecuencia en la que se le presenta cada problema visual y se selecciona la alternativa mediante una escala de Lickert de 5 niveles, que va desde el "Nunca" hasta el "Siempre".

5.1.8 *Visual Performance Questionnaire*

Bergman y Sjöstrand [8] crearon el cuestionario VPQ para medir la prevalencia de la discapacidad visual en los ancianos.

Los autores no consideraron la agrupación de los ITEMS en DOMINIOS, consistiendo este cuestionario en 6 ITEMS únicamente:

1. Ver la televisión.
2. Caminar por el exterior.
3. Lectura de libros y periódicos.
4. Lectura del listín telefónico.
5. Disfrutar de un hobby.
6. Realizar tareas del hogar.

VPQ: VISUAL
PERFORMANCE
QUESTIONNAIRE,
Cuestionario sobre el
Rendimiento Visual.

Las respuestas admisibles son dicotómicas y miden la capacidad de realización de cada tarea por parte de los pacientes.

5.1.9 *Activities of Daily Vision Scale*

ADVS: ACTIVITIES OF DAILY VISION SCALE, Escala de Visión de Actividades Diarias.

La evaluación de la función visual mediante el cuestionario ADVS fue desarrollada por Mangione *et al.* [66].

Este cuestionario consta de 22 ITEMS agrupados en 5 DOMINIOS:

1. Visión a distancia (excluyendo la conducción).
2. Visión cercana.
3. Deslumbramiento.
4. Conducción diurna.
5. Conducción nocturna.

En las respuestas se solicita graduar la dificultad que le requiere a los pacientes la realización de cada una de las actividades indicadas en los ITEMS. Esta graduación se realiza utilizando una escala de Lickert de 5 niveles, que va desde “Sin dificultad” hasta “He dejado de hacerlo debido a la visión”. Posteriormente, se añadió una 6ª opción en la que se indica que el paciente ha dejado de hacer la actividad (o no lo ha hecho nunca) por causas distintas a los problemas de visión.

5.1.10 *Assessment of Visual Function-Related Quality of Life*

VF-QOL: VISUAL FUNCTION-RELATED QUALITY OF LIFE, Calidad de Vida relativa a la Función Visual.

Para medir tanto la función visual como la calidad de vida de los pacientes, Brenner *et al.* [14] desarrollaron el cuestionario VF-QOL.

Este cuestionario consiste en 8 ITEMS:

1. Leer periódicos.
2. Leer el listín telefónico.
3. Leer etiquetas.
4. Leer precios.
5. Reconocer personas.
6. Ver los escalones.
7. Ver grietas en el pavimento.
8. Ver señales de tráfico.

A pesar de que los autores originales no los agruparon en DOMINIOS, Massof *et al.* [68] determinaron, utilizando análisis factorial, que los ITEMS pueden agruparse en 3 DOMINIOS:

1. Visión cercana (ITEMS 1 al 4).
2. Visión de media distancia (ITEMS 5, 6 y 7).
3. Visión lejana (ITEM 8).

Las respuestas a los ITEMS miden la frecuencia con que los pacientes han sentido alguna vez que su visión ha significado un problema para realizar cada ITEM. Cada respuesta tiene 3 posibles opciones: *Nunca*, *A veces*, *Siempre*.

5.1.11 14-Item Visual Functioning Index

Diseñado por Steinberg *et al.* [106] para medir la función visual de pacientes con cataratas, el cuestionario VF-14 requirió un esfuerzo importante, que involucró a muchos expertos de diversas disciplinas (desde oftalmólogos hasta estadísticos, pasando por ópticos y otros médicos especialistas) y también a gran cantidad de pacientes con cataratas (tanto antes como después de la operación de cataratas).

Este cuestionario está diseñado con 14 ITEMS:

1. Lectura de pequeñas impresiones (etiquetas, etc.)
2. Lectura de libros o periódicos.
3. Lectura de libros o periódicos con letras de gran tamaño.
4. Reconocer personas.
5. Ver escalones, escaleras, etc.
6. Reconocer señales de tráfico.
7. Realizar trabajos manuales de precisión: coser, carpintería, etc.
8. Escribir cheques y formularios.
9. Jugar a juegos: bingo, dominó, cartas, ...
10. Practicar deportes.
11. Cocinar.
12. Ver la televisión.
13. Conducción diurna.
14. Conducción nocturna.

Los autores diseñaron el cuestionario basándose en una división de los ITEMS en 4 DOMINIOS, y además determinaron que estadísticamente los datos sustentaban esta división. Sin embargo, en la validación detectaron que pocos pacientes respondían a *todos* los ITEMS, por lo que el conjunto de pacientes no era representativo y no era posible estimar los DOMINIOS. De tal forma que los autores decidieron asumir que los 14 ITEMS formaban parte de un único DOMINIO.

5.1.12 Vision-Related Quality of Life Questionnaire

Frost *et al.* [40] desarrollaron un cuestionario, denominado VCM₁, para evaluar la calidad de vida de los pacientes en relación con su nivel de visión.

Este cuestionario contiene 10 ITEMS integrados todos ellos en un único DOMINIO:

1. Bochorno.
2. Ira.

VF-14: VISUAL FUNCTIONING INDEX WITH 14-ITEM, Índice de Función Visual utilizando 14 ITEMS.

VCM₁: VISION QUALITY-OF-LIFE CORE MEASURE 1, Medida Esencial 1 para determinar la Calidad de Vida según la Visión.

3. Depresión.
4. Soledad.
5. Miedo por el deterioro de la visión.
6. Seguridad en el hogar.
7. Seguridad fuera del hogar.
8. Controlar su vida diaria.
9. Imposibilidad de realizar sus actividades preferidas.
10. Interferencia en su vida.

Las respuestas posibles contestan a dos tipos de preguntas: “cómo su visión ha interferido con alguna determinada actividad”, o “con qué frecuencia la realizan”. Las respuestas están dentro de una escala de 6 opciones graduadas, que se presentan con dos enunciados posibles: el primero, va desde el “No, en absoluto” hasta el “Imposible hacerlo debido a mi visión”; y el segundo, va desde el “Nunca” hasta el “Todo el tiempo”.

5.1.13 *National Eye Institute's Visual Functioning Questionnaire*

NEI: NATIONAL EYE INSTITUTE, Instituto Nacional del Ojo.

NEI-VFQ: NATIONAL EYE INSTITUTE'S VISUAL FUNCTIONING QUESTIONNAIRE, Cuestionario de Funcionamiento Visual del NEI.

El NEI contrató a la compañía Rand Corp. para diseñar el cuestionario NEI-VFQ, para implementar un cuestionario, lo más general posible, para evaluar la calidad de vida en relación con la salud general del paciente y que fuese capaz de evaluar pacientes con un amplio rango de problemas y discapacidades visuales.

El cuestionario tiene 51 ITEMS, que según los autores se asignan a 13 DOMINIOS distintos:

1. Salud general.
2. Visión general.
3. Dolor ocular.
4. Expectativas de visión.
5. Visión cercana.
6. Visión lejana.
7. Problemas sociales.
8. Salud mental.
9. Problemas de rol o modelo de conducta.
10. Dependencia.
11. Conducción.
12. Visión periférica.
13. Visión en color.

Sin embargo, *Massof et al.* [68] realizó un análisis factorial sobre los datos y estos no soportan la existencia de más de 4 DOMINIOS diferentes:

1. Salud general.
2. Dolor ocular.
3. Expectativas de visión.
4. Funcionalidad visual en el día a día.

En el DOMINIO 4, se podrían incluir los 10 DOMINIOS restantes de la propuesta original de los autores.

Las respuestas a los diferentes ITEMS consisten en el grado de dificultad, la frecuencia, el nivel de calidad o el grado de aceptación de cada uno de ellos. Todas las respuestas utilizan una escala de Lickert de 5 puntos que, dependiendo del tipo de pregunta, puede ir desde el “Excelente” al “Pobre”, por ejemplo. Finalmente, se añadió una 6ª opción que está asignada a la respuesta “No lo realizó por otras razones o No estoy interesado en hacer esto”.

5.1.14 Cuestionario de Calidad Visual del Consejo Argentino de Oftalmología

Aparte del trabajo de *Massof et al.* [68], muchas entidades oftalmológicas han realizado cuestionarios para la evaluación de la calidad visual de los pacientes. Muchos de estos cuestionarios son aproximaciones *ad hoc*, sin estudios científicos que los avalen. Sin embargo, fuera del citado estudio de *Massof et al.* cabe destacar el cuestionario **CCV-07-MUDES** propuesto y avalado por el CAO [23], lo cual confiere a este cuestionario un cierto grado de aceptación.

Este cuestionario asistido, es decir, guiado por personas con buena agudeza visual, se basa en la suma de los valores de todas las respuestas, de tal manera que la puntuación de una visión más favorable sería de 37 puntos, mientras que la puntuación más desfavorable sería de 111. Este cuestionario está compuesto por 37 ITEMS agrupados en 8 DOMINIOS:

1. Localización de objetos (*Exacta, Aproximada, No tengo idea*).
2. Percepción de obstáculos (*Exacta, Aproximada, No tengo idea*).
3. Alimentación. (*Correcta, Regular, No lo veo*).
4. Aseo: higiene, vestimenta (*Correcta, Con dificultad, Imposible*).
5. Marcha o movilidad (*Segura, Poco segura, Imposible*).
6. Percepción de objetos móviles (*Correcta, Algunos, Ninguno*).
7. Función visual lejana (*Correcta, Con dificultad, Imposible*).
8. Función visual cercana (*Correcta, Con dificultad, Imposible*).

Este cuestionario está diseñado para proporcionar un valor numérico, que es comparable. En cierto modo, el valor facilita la comprensión del grado de BAJA VISIÓN por parte de los pacientes, lo cual también les permite estimar la evolución de su patología desde un punto de vista menos subjetivo.

CCV-07-MUDES:
CUESTIONARIO DE
CALIDAD VISUAL-07
CON CAPACIDAD
RESIDUAL MENOR A
UNA DÉCIMA EN LA
ESCALA DE SNELLEN
CAO: CONSEJO
ARGENTINO DE
OFTALMOLOGÍA

5.2 DOMINIOS RELEVANTES PARA BAJA VISIÓN Y PROCESAMIENTO DE IMÁGENES

Tras la descripción de los diversos cuestionarios, se ha realizado un estudio de los DOMINIOS utilizados en cada uno de los estudios. Este análisis va a permitir extraer aquellas DIMENSIONES más representativas para el procesamiento de imágenes, con vinculación específica a pacientes con BAJA VISIÓN. Para ello, primero, se han recopilado las DIMENSIONES de todos los cuestionarios analizados en este capítulo, obteniéndose un total de 110 DOMINIOS. Posteriormente, se han unificado todos aquellos DOMINIOS repetidos en más de un cuestionario, lo que ha reducido el tamaño de la lista a 66. Al integrar todos aquellos DOMINIOS que presentan evidentes similitudes, cercanía semántica o que son etimológicamente cercanos, la lista se ha reducido a 22 DOMINIOS. En la Tab. 5.1 se muestran dichas DIMENSIONES. En esta tabla, se han marcado todos aquellos cuestionarios en los que aparecen ITEMS englobados en cada uno de los DOMINIOS estudiados. La última columna de dicha tabla muestra, para cada una de las DIMENSIONES, el número de cuestionarios en los que aparece.

Basándose en dicha tabla, se han eliminado todos aquellos DOMINIOS que no son comportamientos generalizados (es decir, se retiran todos aquellos que no estén en, al menos, 4 cuestionarios diferentes). Con esto se han obtenido un total de 10 DOMINIOS. De estos, se pueden eliminar todos aquellos que no traten directamente con funciones de tipo visual, por ejemplo, se pueden eliminar todos los DOMINIOS que traten de la salud mental, problemas sociales, etc., de los pacientes, como el D22 (*Otros rasgos no directamente visuales: salud, dolor, problemas sociales*). También es posible retirar todas aquellas DIMENSIONES que queden fuera del ámbito de actuación efectivo de los algoritmos de extracción de bordes, como pueden ser los que involucran imágenes en movimiento, ej. D5 (*Ver la televisión*), D6 (*Conducir*) o D7 (*Otras actividades: ocio, ir de compras, jugar juegos mesa*). Hay que indicar que todos los cuestionarios se encuentran representados al menos en uno de los DOMINIOS seleccionados.

Para finalizar, se han preseleccionado 7 DOMINIOS para el desarrollo de la encuesta de evaluación de contornos vinculados a pacientes con BAJA VISIÓN. En esta preselección, realizada con la metodología descrita, se han escogido los DOMINIOS que en la Tab. 5.1 han obtenido un mayor número de ocurrencias en los diferentes cuestionarios y se han rechazado aquellos grupos que están fuera del ámbito de trabajo de esta Tesis Doctoral.

Los DOMINIOS preseleccionados han sido los siguientes:

- D1 *Visión media/larga distancia, movilidad general, orientación y seguridad en exteriores, ver grietas en pavimento y ver señales de tráfico.*
- D2 *Agudeza visual en la visión cercana, lectura de libros, periódicos, listín telefónico, escritura de cheques.*
- D3 *Orientación y seguridad en interiores.*

ID	DOMINIO	VFI	RAND	VF-PRP	VSI	LVF	FALV	V AQ	VTQ	ADYS	VF-QOL	VF-14	VCM1	NEI-VTQ	CCV07	N
D1	Visión media/larga distancia, movilidad general, orientación y seguridad en exteriores, ver grietas en pavimento y ver señales de tráfico	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	11
D2	Agudeza visual en la visión cercana, lectura de libros, periódicos, listín telefónico, escritura de cheques	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	11
D3	Orientación y seguridad en interiores	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
D4	Sensibilidad y adaptación a la luz y problemas de iluminación, deslumbramientos, y visión y conducción nocturna	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
D5	Ver la televisión	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4
D6	Conducción diurna	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4
D7	Otras actividades: ocio, ir de compras, jugar juegos mesa	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4
D8	Cuidado personal: aseo, baño, comer	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4
D9	Reconocer personas	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4
D10	Actividad principal: trabajo, tareas del hogar	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	3
D11	Visión binocular	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	3
D12	Color	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	3
D13	Velocidad de procesamiento visual	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	2
D14	Visión periférica	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	2
D15	Viajar en autobús	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D16	Falta de fluidez visual	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D17	Búsqueda	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D18	Detección	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D19	Análisis	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D20	Seguimiento	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D21	Practicar deportes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
D22	Otros rasgos no directamente visuales: salud, dolor, problemas sociales	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	4

Tabla 5.1
Determinación de DOMINIOS para el procesamiento de imágenes y BAJA VISIÓN.
N: Número de cuestionarios en los que aparece dicho DOMINIO.

- D4 *Sensibilidad y adaptación a la luz y problemas de iluminación, deslumbramientos y Visión y conducción nocturna.*
- D8 *Cuidado personal: aseo, baño, comer.*
- D9 *Reconocer personas.*
- D10 *Actividad principal: trabajo y tareas del hogar.*

Esta preselección permite establecer unas bases mínimas de la tipología de imágenes que se deben buscar para realizar la encuesta de evaluación de calidad de contornos, para que dicha encuesta tenga utilidad para pacientes con BAJA VISIÓN. De entre estos 7 DIMENSIONES preseleccionadas, en el Cap. 7, se escogerán 5 DOMINIOS, en función de su adecuación con los objetivos de la Tesis.

5.3 CONCLUSIONES

Para el diseño de cuestionarios que involucren a pacientes con BAJA VISIÓN es conveniente conocer qué aspectos se preguntan en otros cuestionarios para dicho tipo de pacientes. En este capítulo se han descrito los cuestionarios más importantes en este ámbito. También, se ha realizado un análisis de los DOMINIOS más relevantes, obteniendo finalmente una preselección de 7 DIMENSIONES.

Así pues, como colofón de estos capítulos introductorios, este capítulo ha permitido completar la exposición de contenidos previos necesarios para construir experimentos que sean útiles para pacientes con BAJA VISIÓN y para diseñar cuestionarios que permitan evaluar la calidad de los métodos con los que se generan los experimentos. Estos cuestionarios servirán para que pacientes con BAJA VISIÓN realicen la evaluación de la calidad subjetiva de los experimentos, aunque también permitirán comparar las respuestas de otros sujetos que no sufran esta patología ante los mismos estímulos.

5.3.1 Conclusiones de la Parte I

Con este capítulo se finaliza la Parte I, *Introducción*, de la Tesis Doctoral. En los siguientes capítulos, se proponen las principales aportaciones científicas propias de esta Tesis, tanto en el ámbito del PROCESAMIENTO DE IMÁGENES para la obtención de contornos en imágenes con iluminación no homogénea como en el diseño de cuestionarios para la evaluación subjetiva de la calidad de los algoritmos propuestos, que se engloban en la Parte II, *Método y Materiales*.

Parte II

MÉTODO Y MATERIALES

ÍNDICE

6.1	La convolución en el procesamiento de imágenes	104
6.1.1	Descripción matemática de la convolución	104
6.1.2	Convolución 2D con filtros separables	105
6.2	Convolución LIP 2D con filtros separables	106
6.2.1	Desarrollo matemático de la Convolución LIP 2D	106
6.2.2	Convolución LIP Generalizada Directa: DGLIP-Conv	108
6.2.3	Convolución LIP Generalizada Rápida: FGLIP-Conv	109
6.3	Rango de funcionamiento de la Convolución LIP	109
6.3.1	Límite inferior del rango	109
6.3.2	Límite superior del rango	110
6.3.3	Límite inferior y superior del rango	111
6.4	Comparativas de rendimiento	111
6.4.1	Coste computacional por operador	113
6.5	Conclusiones	114

El modelo LIP presenta muchas características beneficiosas en el procesamiento de imágenes naturales, como por ejemplo, procesado homogéneo en zonas con iluminación no uniforme, alta sensibilidad en zonas oscuras, robustez matemática, etc. Para aprovechar estas ventajas, algunos algoritmos tradicionales se han portado al paradigma LIP mediante aproximaciones *ad hoc*. En muchos casos, las implementaciones de los algoritmos bajo LIP son tan específicas que obligan a recodificar incluso tras un cambio de parámetros, lo que ha limitado en gran medida el avance de este paradigma.

Una gran cantidad de algoritmos de procesamiento de imágenes utilizan la *convolución* como operador principal, lo que hace que dicho operador se considere entre los más importantes dentro de este campo. Por ello, en este capítulo se presentan las bases matemáticas de un nuevo operador, *Convolución-LIP*, que se propone como una de las principales aportaciones de esta Tesis Doctoral. Esta adaptación se ha diseñado para sustituir al operador tradicional *convolución* bajo el modelo LIP, y para que, además, se puedan aplicar de manera simple y eficaz, los filtros de procesamiento de imágenes utilizando este modelo. Este nuevo operador permite implementar fácilmente una gran cantidad de algoritmos de procesamiento de imágenes. Esta aportación científica se describe matemáticamente en este capítulo, indicando rangos de funcionamiento y se presentan dos métodos diferentes para la implementación de dicho operador.

6.1 LA CONVOLUCIÓN EN EL PROCESAMIENTO DE IMÁGENES

Muchos algoritmos de procesamiento de imágenes y vídeo utilizan el operador *convolución* como uno de los primeros componentes matemáticos en el proceso de obtención del resultado. Esto provoca, a su vez, que la *convolución* sea, de manera generalizada, uno de los elementos principales para el desarrollo de los procedimientos de procesado de imágenes y vídeo.

También hay que indicar que la *convolución* es un operador con alto consumo computacional. Este alto consumo provoca que para muchos algoritmos de procesamiento de imágenes, el porcentaje de tiempo de cómputo consumido por la *convolución* sea muy elevado, en algunos casos superior al 50% del tiempo total del algoritmo. Por ello, se han realizado muchos esfuerzos por mejorar el rendimiento de este operador, tanto a nivel matemático como a nivel computacional, por ejemplo, proponiendo implementaciones paralelas eficientes.

Aunque en la mejora del operador *convolución* no sólo debe tenerse en cuenta el aspecto del rendimiento temporal, sino también si la implementación propuesta permite realizar otras operaciones adicionales. Por ejemplo, una determinada implementación de la *convolución* será mejor que la implementación básica habitual, si además de la operación *convolución* tradicional, que es la que única que se aplica habitualmente en la implementación básica, es capaz de aplicar una función de transformación del espacio de trabajo. En particular, este aspecto es uno de los más beneficiosos que permite el nuevo operador que se propone en este capítulo.

6.1.1 Descripción matemática de la convolución

La *convolución* es un operador matemático que aplica un determinado filtro sobre una entrada, en las que tanto la imagen como el filtro pueden ser discretos o continuos, y también unidimensionales, bidimensionales o multidimensionales. De manera habitual, el operador *convolución* utilizado en el ámbito del procesamiento de imágenes opera con filtros discretos unidimensionales o bidimensionales y con imágenes discretas bidimensionales, y es este tipo de operador el que se va a considerar en esta Tesis.

Los filtros serán, respectivamente, matrices bidimensionales o vectores unidimensionales compuestos por valores escalares discretos. A continuación, se muestra la formulación matemática de la *convolución* con filtros bidimensionales:

$$\begin{aligned}
 C = \text{conv2D}(I, F) &= \begin{pmatrix} F(1,1) & \dots & F(1, n_f) \\ \vdots & \ddots & \vdots \\ F(m_f, 1) & \dots & F(m_f, n_f) \end{pmatrix} \otimes \begin{pmatrix} I(1,1) & \dots & I(1, w) \\ \vdots & \ddots & \vdots \\ I(h, 1) & \dots & I(h, w) \end{pmatrix} \\
 &= \begin{pmatrix} C(1,1) & \dots & C(1, w) \\ \vdots & \ddots & \vdots \\ C(h, 1) & \dots & C(h, w) \end{pmatrix}
 \end{aligned}$$

Donde F representa a un filtro bidimensional de tamaño $m_f \times n_f$ elementos e I representa una imagen de tamaño $h \times w$.

$$C(i, j) = \sum_{m=0}^{m_f-1} \sum_{n=0}^{n_f-1} (F(m, n) \cdot I(i-m, j-n)) \quad (6.1)$$

con $1 \leq i \leq h$, $1 \leq j \leq w$.

6.1.2 Convolución 2D con filtros separables

Si la matriz bidimensional que compone el filtro se puede construir a partir de la multiplicación matricial de dos vectores unidimensionales, se dice que dicho filtro es separable. La separación de los filtros permite que una convolución con un filtro bidimensional, se pueda realizar mediante dos convoluciones consecutivas con filtros unidimensionales. La gran mayoría de algoritmos de procesamiento de imágenes utilizan filtros que son separables. Por tanto, en esta Tesis Doctoral se va a trabajar utilizando el operador *convolución* con filtros bidimensionales separables (por abreviar, *convolución 2D*), realizada a partir de dos *convoluciones 1D* (denominación abreviada de *convolución* que usa filtros unidimensionales) consecutivas. A continuación se expone la formulación matemática que describe este tipo de *convolución*:

$$\text{conv2D}(I, F) = \overbrace{\text{conv1D}(\underbrace{\text{conv1D}(I, \mathbf{a}^T)}_{\mathbf{c}_a}, \mathbf{b})}_{\mathbf{c}_b} \quad (6.2)$$

en la que I representa a una imagen y F es el filtro que se desea aplicar espacialmente. En este caso, el filtro F es una matriz bidimensional que se puede separar en dos vectores fila unidimensionales \mathbf{a} , \mathbf{b} , tal que $F = \mathbf{a}^T \times \mathbf{b}$. Donde \mathbf{a}^T representa la transpuesta del vector \mathbf{a} , es decir, un vector columna. Para simplificar el proceso se ha tomado que el tamaño del filtro coincide con el tamaño de la imagen, aspecto este que no afecta a la formulación matemática en absoluto, salvo por la claridad en el desarrollo.

La ecuación (6.2) se puede desarrollar, de manera sencilla, expandiendo sus términos. A continuación, se describe el contenido de \mathbf{c}_a , que es el resultado de la *convolución 1D* de la imagen por el vector \mathbf{a}^T :

$$\begin{aligned} \mathbf{c}_a &= \begin{pmatrix} \mathbf{a}(1) \\ \vdots \\ \mathbf{a}(m) \end{pmatrix} \otimes \begin{pmatrix} I(1,1) & \dots & I(1,n) \\ \vdots & \ddots & \vdots \\ I(m,1) & \dots & I(m,n) \end{pmatrix} \\ &= \left[\mathbf{a}(m) \cdot I(1,1) + \dots + \mathbf{a}(1) \cdot I(m,1), \right. \\ &\quad \mathbf{a}(m) \cdot I(1,2) + \dots + \mathbf{a}(1) \cdot I(m,2), \dots, \\ &\quad \left. \mathbf{a}(m) \cdot I(1,n) + \dots + \mathbf{a}(1) \cdot I(m,n) \right] \\ &= \text{vector}_{i=0}^{n-1} \left[\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot I(i+1, j+1) \right] = [\mathbf{c}_a(1), \dots, \mathbf{c}_a(n)] \end{aligned}$$

Donde $\mathbf{a}(i)$ representa el valor i -ésimo del vector \mathbf{a} e $I(i, j)$ es el valor de la imagen que se encuentra en la columna j -ésima de la fila i -ésima. De manera análoga, siguiendo la misma descripción anterior, se puede obtener $\mathbf{c}_b = \text{conv1D}(\mathbf{c}_a, \mathbf{b})$, como:

$$\begin{aligned} \mathbf{c}_b &= \mathbf{b}(n) \cdot \mathbf{c}_a(1) + \dots + \mathbf{b}(1) \cdot \mathbf{c}_a(n) = \sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \mathbf{c}_a(i+1) \\ &= \sum_{i=0}^{n-1} \left(\mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot I(i+1, j+1) \right) \right) \end{aligned}$$

Con lo que se obtiene la fórmula conocida de la *convolución* 2D a partir de dos *convoluciones* 1D:

$$\text{conv2D}(I, F) = \sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot I(i+1, j+1) \right) \quad (6.3)$$

Esta formulación de la *convolución* 2D con filtros separables permite reducir la complejidad de los cálculos frente a la *convolución* 2D con filtros bidimensionales, ya que se utilizan un menor número de iteraciones para calcular el mismo resultado. Esto reduce la complejidad computacional de $O(n^4)$ en la implementación de la *convolución* 2D con filtros bidimensionales a $2 \cdot O(n^3)$ para la *convolución* 2D con filtros separables, por lo que, siempre que sea posible es preferible utilizar este último método.

6.2 CONVOLUCIÓN LIP 2D CON FILTROS SEPARABLES

En esta sección se expondrá toda la formulación matemática necesaria para calcular la versión LIP de la *convolución* de (6.3), a la que se llamará *Convolución-LIP* [80, 82]. Como se ha comentado en la Sec. 2.2, el paradigma LIP permite adaptar las fórmulas del procesamiento clásico de imágenes, de dos maneras diferentes: mediante un método “directo” o mediante un método basado en una “transformada”. En este caso, se obtendrá la *Convolución-LIP* mediante el método “directo”, en la que se tendrá una fórmula con operadores modificados LIP y con la imagen en *tonos de gris*. Una vez se tenga una fórmula que utilice operadores LIP, se desarrollará esa fórmula para que en la misma sólo se encuentren operadores tradicionales.

6.2.1 Desarrollo matemático de la *Convolución LIP 2D*

Así pues, tomando como inicio (6.3), se aplica el mecanismo “directo” del paradigma LIP, sustituyendo los operadores tradicionales por los operadores modificados LIP (Δ , $\hat{\Delta}$) y utilizando la imagen en *tono de gris*, \hat{I} . Quedando la siguiente ecuación:

$$\text{conv2D}_{\text{LIP}}(\hat{I}, F) = \sum_{i=0}^{n-1} \Delta \mathbf{b}(n-i) \Delta \left(\sum_{j=0}^{m-1} \Delta \mathbf{a}(m-j) \Delta \hat{I}(i+1, j+1) \right) \quad (6.4)$$

Para utilizar sólo operadores tradicionales, se aplica en ambas partes la transformada isomórfica, (2.16). Como los valores de los vectores filtro son valores escalares, no es necesario aplicarles ninguna función de transformación, por tanto, la transformada isomórfica se aplica únicamente a la imagen.

$$\varphi(\text{conv2D}_{\text{LIP}}(\widehat{I}, F)) = \sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \varphi(\widehat{I}(i+1, j+1)) \right) \quad (6.5)$$

Tomando (6.5), se desarrolla la función transformada isomórfica, se renombra $(M - \widehat{I}(i+1, j+1))$ como $I(i+1, j+1)$ y se continúa operando:

$$\begin{aligned} \varphi(\text{conv2D}_{\text{LIP}}(\widehat{I}, F)) &= \\ &= \sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \left(-M \cdot \ln \left(\frac{M - \widehat{I}(i+1, j+1)}{M} \right) \right) \right) \\ &= -M \cdot \sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \left(\ln \left(\frac{I(i+1, j+1)}{M} \right) \right) \right) \\ &= -M \cdot \sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left[\left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \ln(I(i+1, j+1)) \right) \right. \\ &\quad \left. - \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \ln(M) \right) \right] \\ &= -M \cdot \left[\sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \ln(I(i+1, j+1)) \right) \right. \\ &\quad \left. - \left(\ln(M) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \sum_{i=0}^{n-1} \mathbf{b}(n-i) \right) \right) \right] \quad (6.6) \end{aligned}$$

Para simplificar la ecuación anterior, se puede definir la siguiente constante:

$$K = \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \sum_{i=0}^{n-1} \mathbf{b}(n-i) \right) \quad (6.7)$$

Gracias a la que se obtiene (6.8).

$$\begin{aligned} \varphi(\text{conv2D}_{\text{LIP}}(\widehat{I}, F)) &= \\ &= -M \cdot \left[\sum_{i=0}^{n-1} \mathbf{b}(n-i) \cdot \left(\sum_{j=0}^{m-1} \mathbf{a}(m-j) \cdot \ln(I(i+1, j+1)) \right) - K \cdot \ln(M) \right] \quad (6.8) \end{aligned}$$

6.2.2 Convolución LIP Generalizada Directa: DGLIP-Conv

En la sección anterior se ha descrito el operador *Convolución*-LIP desde un punto de vista estrictamente matemático. En esta sección se describe una implementación específica del nuevo operador propuesto, que se ha llamado **DGLIP-CONV**. Se ha denominado este método como *Directo* puesto que se usa el mecanismo directo de aplicación de LIP: uso de imágenes "originales" con operadores especiales.

Tomando como base (6.8) y sabiendo que $a \cdot \ln x = \ln x^a$, se puede obtener la siguiente reformulación:

$$\begin{aligned} \varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) = \\ -M \cdot \left[\sum_{i=0}^{n-1} \mathbf{b}^{(n-i)} \cdot \left(\sum_{j=0}^{m-1} \ln \left(I_{(i+1, j+1)}^{\alpha(m-j)} \right) \right) - K \cdot \ln(M) \right] \end{aligned} \quad (6.9)$$

Como se sabe, $\sum_i (\ln x_i) = \ln \left(\prod_i x_i \right)$, y por tanto:

$$\begin{aligned} \varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) = \\ -M \cdot \left[\sum_{i=0}^{n-1} \mathbf{b}^{(n-i)} \cdot \ln \left(\prod_{j=0}^{m-1} I_{(i+1, j+1)}^{\alpha(m-j)} \right) - K \cdot \ln(M) \right] \end{aligned}$$

Se vuelve a aplicar a la fórmula el hecho de que $a \cdot \ln x = \ln x^a$, obteniéndose:

$$\begin{aligned} \varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) = \\ -M \cdot \left[\sum_{i=0}^{n-1} \ln \left(\left(\prod_{j=0}^{m-1} I_{(i+1, j+1)}^{\alpha(m-j)} \right)^{\mathbf{b}^{(n-i)}} \right) - K \cdot \ln(M) \right] \end{aligned}$$

Finalmente, se reordena la fórmula retirando el valor negativo del factor M que multiplica a toda la fórmula y se obtiene:

$$\begin{aligned} \varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) = \\ M \cdot \left[K \cdot \ln(M) - \ln \left(\prod_{i=0}^{n-1} \left(\prod_{j=0}^{m-1} I_{(i+1, j+1)}^{\alpha(m-j)} \right)^{\mathbf{b}^{(n-i)}} \right) \right] \end{aligned} \quad (6.10)$$

Por último, aplicando la transformada isomórfica inversa (2.17) a (6.10), se obtiene:

$$\begin{aligned} \text{conv2D}_{\text{LIP}}(\hat{I}, F) = \varphi^{-1} \left(\varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) \right) = \\ M \cdot \left(1 - \frac{\prod_{i=0}^{n-1} \prod_{j=0}^{m-1} \left(I_{(i+1, j+1)}^{\alpha(m-j)} \right)^{\mathbf{b}^{(n-i)}}}{M^K} \right) \end{aligned} \quad (6.11)$$

6.2.3 Convolución LIP Generalizada Rápida: FGLIP-Conv

La formulación que se ha presentado en la sección anterior (DGLIP-CONV) utiliza los operadores división y exponenciación, que son computacionalmente costosos. Para evitar el uso de estos operadores, se ha buscado otra formulación alternativa a la presentada anteriormente que utilice operaciones menos costosas. Esta nueva formulación sigue otra aproximación diferente para el procesado LIP (es decir, aplica los operadores usuales a imágenes “transformadas”). Con esta nueva formulación, además de obtenerse los mismos resultados que con DGLIP-CONV, se reduce el tiempo necesario para calcular el resultado, ya que se usan operadores computacionalmente más simples. Por este motivo a la convolución obtenida mediante este método se la ha denominado *Rápida*. Por tanto, para utilizar una nomenclatura análoga a la del método propuesto con anterioridad, este método se ha llamado **FGLIP-CONV**.

FGLIP-CONV: Del inglés, FAST GENERAL LIP-CONVOLUTION, Convolución-LIP Generalizada Rápida

Utilizando K , definida en (6.7), se puede re-escribir la ecuación (6.8) como:

$$\varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) = M \cdot \left(K \cdot \ln(M) - \text{conv1D} \left(\text{conv1D}(\ln I, \mathbf{a}^T), \mathbf{b} \right) \right) \quad (6.12)$$

Análogamente, si se aplica la transformación isomórfica inversa (2.17) sobre 6.12, se puede obtener:

$$\text{conv2D}_{\text{LIP}}(\hat{I}, F) = \varphi^{-1} \left(\varphi(\text{conv2D}_{\text{LIP}}(\hat{I}, F)) \right) = M \cdot \left(1 - e^{\text{conv1D}(\text{conv1D}(\ln I, \mathbf{a}^T), \mathbf{b}) - K \cdot \ln(M)} \right) \quad (6.13)$$

6.3 RANGO DE FUNCIONAMIENTO DE LA CONVOLUCIÓN LIP

Como cualquier operador LIP (excepto la resta LIP y la multiplicación LIP con valores escalares negativos) el resultado de la *Convolución-LIP* se restringe al intervalo $(0, M)$.

Con las siguientes operaciones matemáticas, se demuestra que la aplicación del operador *Convolución-LIP* genera resultados dentro del rango de operación de LIP y por tanto, estos resultados pueden ser utilizados como entrada de nuevos operadores LIP. Es decir, en los siguientes apartados se demostrará la siguiente afirmación matemática:

$$\text{conv2D}_{\text{LIP}}(\hat{I}, F) \in (0, M) \subseteq \mathbb{R}$$

6.3.1 Límite inferior del rango

Primero se demuestra que la aplicación del operador *Convolución-LIP* proporciona resultados que, en todo caso, son siempre estrictamente positivos. Para confirmarlo, se utilizará la *reducción al absurdo*. Para lo cual, se niega la afirmación que se desea demostrar y se opera:

$$\text{conv2D}_{\text{LIP}}(\hat{I}, F) \not\geq 0 \Rightarrow \text{conv2D}_{\text{LIP}}(\hat{I}, F) \leq 0$$

$$\begin{aligned}
& M \cdot \left(1 - \frac{\prod_{i=0}^{n-1} \left(\prod_{j=0}^{m-1} I(i+1, j+1) a^{(m-j)} \right)^{b(n-i)}}{M \sum_{j=0}^{m-1} a^{(m-j)} \cdot \sum_{i=0}^{n-1} b(n-i)} \right) \leq 0 \\
& \frac{\prod_{i=0}^{n-1} \left(\prod_{j=0}^{m-1} I(i+1, j+1) a^{(m-j)} \right)^{b(n-i)}}{M \sum_{j=0}^{m-1} a^{(m-j)} \cdot \sum_{i=0}^{n-1} b(n-i)} \geq 1 \\
& \prod_{i=0}^{n-1} \left(\prod_{j=0}^{m-1} I(i+1, j+1) a^{(m-j)} \right)^{b(n-i)} \geq M \sum_{j=0}^{m-1} a^{(m-j)} \cdot \sum_{i=0}^{n-1} b(n-i) \\
& \prod_{i=0}^{n-1} \left(\prod_{j=0}^{m-1} I(i+1, j+1) a^{(m-j)} \right)^{b(n-i)} \geq \prod_{i=0}^{n-1} \left(\prod_{j=0}^{m-1} M a^{(m-j)} \right)^{b(n-i)}
\end{aligned}$$

Lo que implicaría que:

$$\exists(x, y) : I(x, y) \geq M$$

Sin embargo, esto es imposible ya que, por definición, $I \in (0, M) \subseteq \mathbb{R}$. Con lo cual, se puede afirmar con total seguridad que la hipótesis de partida es falsa:

$$\text{conv}2D_{\text{LIP}}(\hat{I}, F) \not\leq 0$$

y por tanto, queda demostrado que:

$$\text{conv}2D_{\text{LIP}}(\hat{I}, F) > 0 \tag{6.14}$$

6.3.2 Límite superior del rango

De manera análoga, también se demostrará mediante *reducción al absurdo* que todos los resultados deben ser estrictamente menores que M .

$$\text{conv}2D_{\text{LIP}}(\hat{I}, F) \not\leq M \Rightarrow \text{conv}2D_{\text{LIP}}(\hat{I}, F) \geq M$$

$$\begin{aligned}
M \cdot \left(1 - \frac{\prod_{i=0}^{n-1} \prod_{j=0}^{m-1} \left(I(i+1, j+1) \mathbf{a}^{(m-j)} \right)^{\mathbf{b}^{(n-i)}}}{M \sum_{j=0}^{m-1} \mathbf{a}^{(m-j)} \cdot \sum_{i=0}^{n-1} \mathbf{b}^{(n-i)}} \right) &\geq M \\
1 - \frac{\prod_{i=0}^{n-1} \prod_{j=0}^{m-1} \left(I(i+1, j+1) \mathbf{a}^{(m-j)} \right)^{\mathbf{b}^{(n-i)}}}{M \sum_{j=0}^{m-1} \mathbf{a}^{(m-j)} \cdot \sum_{i=0}^{n-1} \mathbf{b}^{(n-i)}} &\geq 1 \\
\frac{\prod_{i=0}^{n-1} \prod_{j=0}^{m-1} \left(I(i+1, j+1) \mathbf{a}^{(m-j)} \right)^{\mathbf{b}^{(n-i)}}}{M \sum_{j=0}^{m-1} \mathbf{a}^{(m-j)} \cdot \sum_{i=0}^{n-1} \mathbf{b}^{(n-i)}} &\leq 0 \\
\prod_{i=0}^{n-1} \prod_{j=0}^{m-1} \left(I(i+1, j+1) \mathbf{a}^{(m-j)} \right)^{\mathbf{b}^{(n-i)}} &\leq 0
\end{aligned}$$

Lo que implicaría que:

$$\exists(x, y) : I(x, y) \leq 0$$

Aspecto que es imposible ya que, por definición, $I \in (0, M) \subseteq \mathbb{R}$. Esto implica que la hipótesis de partida es falsa:

$$\text{conv2D}_{\text{LIP}}(\hat{I}, F) \not\geq 0$$

y por tanto, queda demostrado que:

$$\text{conv2D}_{\text{LIP}}(\hat{I}, F) < M \tag{6.15}$$

6.3.3 Límite inferior y superior del rango

Unificando (6.14) y (6.15), se obtiene:

$$0 < \text{conv2D}_{\text{LIP}}(\hat{I}, F) < M$$

Lo que demuestra que el rango de trabajo del operador *Convolución–LIP* es equivalente al de las imágenes *tonos de gris*, con las que trabaja LIP. Por tanto, siempre y cuando el filtro esté compuesto por números reales y la imagen sea de *tonos de gris*, este operador no va a producir resultados “fuera de rango” en ningún caso.

6.4 COMPARATIVAS DE RENDIMIENTO

Para comprobar que, en las mismas condiciones, FGLIP–CONV es más rápido que DGLIP–CONV, se han realizado varios experimentos. Todos los experimentos se han aplicado sobre una imagen de tamaño 1024×1024 píxeles. Sobre esta imagen se han ido aplicando filtros de distintos tamaños, variando desde 3 hasta 515 elementos de longitud. Como para estos experimentos lo que se pretende estimar es el tiempo

de cómputo, los resultados visuales no son necesarios, y por tanto, el contenido de los filtros es irrelevante. Tan sólo tiene influencia en el tiempo de procesamiento el tamaño de los filtros y no su contenido. El experimento muestra la comparación de tiempos de procesamiento para las dos implementaciones (DGLIP-CONV y FGLIP-CONV) con distintos tamaños de filtro. La *convolución* ha sido aplicada mediante dos filtros vector, el primero para la *convolución* horizontal y el segundo para la *convolución* vertical. Los resultados temporales de este experimento se muestran en la Fig. 6.1, mientras que en la Fig. 6.2 se muestra la evolución de la ganancia de velocidad de FGLIP-CONV con respecto a a DGLIP-CONV.

Este experimento se ha ejecutado sobre la plataforma Matlab 7.6.0 (R2008a) bajo Linux 2.6.29 (Fedora 10) en un computador Centrino 2 Duo a 2.2 Ghz con doble núcleo de ejecución, aunque Matlab está configurado para hacer uso de un único núcleo. Tanto para FGLIP-CONV como para DGLIP-CONV se ha implementado la función minimizando el número de bucles y maximizando el número de operaciones ejecutadas sobre matrices frente a las ejecuciones frente a escalares, tal y como recomienda Matlab para optimizar el rendimiento de los programas en su plataforma. Adicionalmente, las dos implementaciones se han compilado para acelerar su ejecución. Esta optimización mejora el tiempo de proceso al eliminar la necesidad de compilación de las instrucciones cíclicas de los bucles.

Por otra parte, como FGLIP-CONV utiliza la función *convolución* estándar, este operador se puede ejecutar utilizando la función *convolución* optimizada que proporciona Matlab en su librería de imágenes en lugar de realizar los cálculos explícitamente. Por ello, se ha construido una versión optimizada de *FGLIP-Conv*, que se ha denominado FGLIP-CONV OPTIM., en la que se hace uso de la función *imfilter* de Matlab para implementar la *convolución*. A pesar de que esta implementación está catalogada como *Optimizada*, según la documentación de Matlab, la función *imfilter* no está optimizada para utilizar las extensiones de Intel para los tipos de datos utilizados en los experimentos. Por tanto, otras implementaciones de *imfilter* mejoradas podrían proporcionar resultados temporales aún mejores.

Como se puede observar en la Fig. 6.1, para todos los tamaños de filtros, FGLIP-CONV es mucho más rápido que DGLIP-CONV. También se puede apreciar en dicha gráfica que la versión optimizada, FGLIP-CONV OPTIM. es mucho más rápida que las otras dos implementaciones.

En la Fig. 6.2, se muestra la ganancia de velocidad de FGLIP-CONV frente a DGLIP-CONV. La ganancia de la versión no optimizada, *FGLIP-Conv*, se encuentra en torno al 2 % de media. El aumento de la ganancia de velocidad es mucho mayor cuando se utiliza la versión *Optimizada*, estando siempre por encima del 10 % y llegando a picos del 37 %, aunque este rendimiento va degradándose poco a poco conforme aumenta el tamaño del filtro. Por el contrario, FGLIP-CONV presenta un aumento suave pero progresivo en la ganancia de velocidad con respecto a los tiempos de DGLIP-CONV.

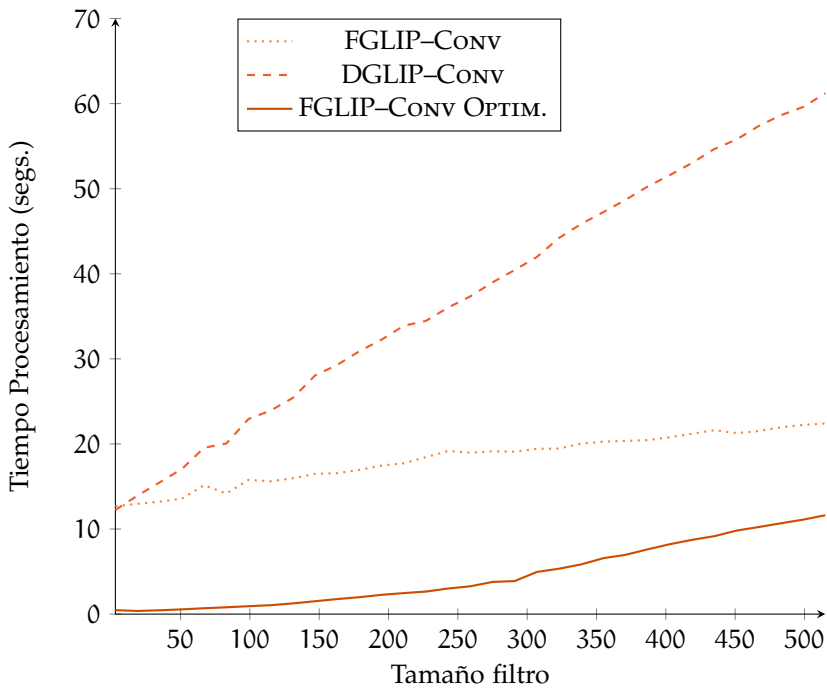


Figura 6.1 Comparativa de tiempos de FGLIP-CONV y DGLIP-CONV para distintos tamaños de filtros sobre una imagen de 1024×1024 píxeles.

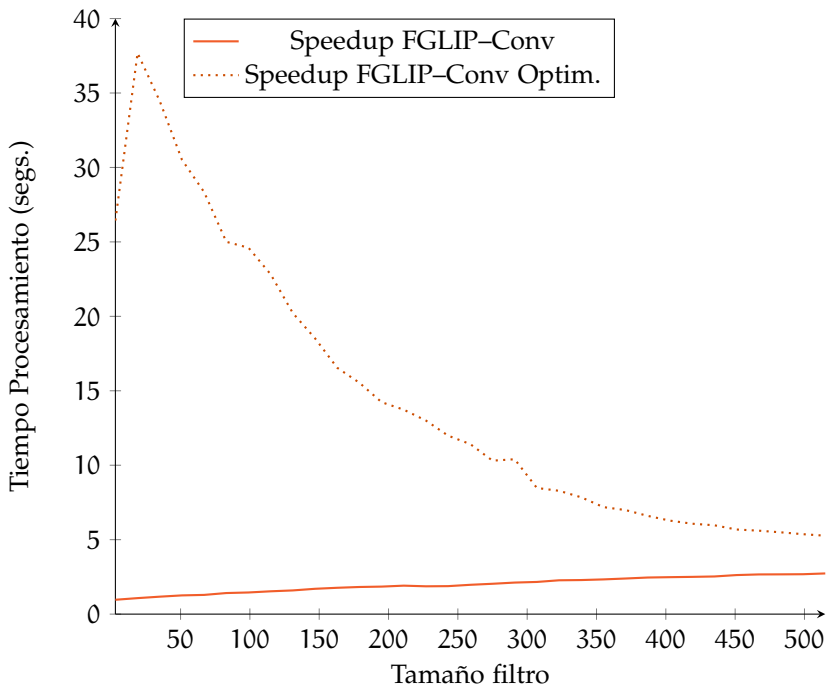


Figura 6.2 Ganancia de velocidad de FGLIP-CONV/FGLIP-CONV OPTIM. frente a DGLIP-CONV para distintos tamaños de filtros sobre la misma imagen de 1024×1024 píxeles.

6.4.1 Coste computacional por operador

Para una misma ejecución particular, se ha realizado un perfilado (*profiling*) de cada línea de código para cada una de las dos implemen-

taciones descritas, obteniéndose los resultados que se muestran en la Tab. 6.1. En ella, se puede observar cómo para la implementación de DGLIP-CONV, el tiempo empleado en la exponenciación es prácticamente igual a la suma del tiempo de todos los operadores utilizados en FGLIP-CONV. Esto demuestra que el coste computacional debido a los operadores es mucho mayor en DGLIP-CONV

Tabla 6.1
Coste computacional
por operador.

Implementación	Operador	Tiempo
FGLIP-CONV	Multiplicación	12.40
	Suma	13.82
DGLIP-CONV	Multiplicación	15.77
	Exponenciación	22.07

6.5 CONCLUSIONES

Este capítulo, en el que se ha presentado toda la formulación matemática para el desarrollo del operador *Convolución-LIP*, constituye el núcleo principal de esta Tesis Doctoral. Este operador permite, de manera sencilla y homogénea, adaptar una gran cantidad de algoritmos de procesamiento de imágenes y vídeo al paradigma LIP. Como resumen se muestran a continuación las principales aportaciones científicas obtenidas a partir de la investigación desarrollada en este capítulo:

- *Convolución-LIP*: Se ha propuesto este nuevo operador diseñado matemáticamente utilizando la fortaleza y robustez que proporciona la estructura algebraica del paradigma LIP. Este diseño matemático ha permitido construir dos implementaciones diferentes:
 - A. DGLIP-CONV: Esta implementación se obtiene a través del método “directo”, utilizando las imágenes originales sin transformar y un conjunto de operadores matemáticos modificados. Tras operar matemáticamente se consigue una fórmula que utiliza multiplicaciones y exponenciaciones.
 - B. FGLIP-CONV: Parte del método “transformado”, que utiliza las imágenes aplicándoles una función de transformación y los operadores matemáticos clásicos, multiplicación y suma.

Hay que indicar que ambas implementaciones son equivalentes en términos de resultados, gracias a lo cual se tienen dos mecanismos diferentes para aplicar uno de los operadores más extendidos en el procesamiento de imágenes, como es la *convolución* bajo un modelo LIP. Éste posee unos beneficios muy importantes al trabajar con imágenes: coherencia de las operaciones con leyes físicas y psico-físicas, todas las operaciones LIP se mantienen dentro del rango de trabajo, los algoritmos presentan una cierta inmunidad ante efectos de iluminación no homogénea y muestran una mayor sensibilidad en zonas oscuras. A pesar de la equivalencia de las implementaciones, unos experimentos han

mostrado que la implementación FGLIP-CONV calcula la *Convolución-LIP* de manera más rápida que la implementación DGLIP-CONV.

La *convolución* es un operador que suele estar implementado y optimizado en diferentes arquitecturas de computación, lo cual proporciona un mayor rendimiento que las implementaciones estrictamente software de dicho operador. Gracias a que FGLIP-CONV permite utilizar la *convolución* estándar, se ha mostrado que es posible construir dicho operador haciendo uso de la *convolución* optimizada presente en la mayoría de sistemas de computación, con un rendimiento aún más eficiente que la versión que implementa la *convolución* a partir de sus operadores básicos.

En este capítulo se ha demostrado que el nuevo operador *Convolución-LIP* proporciona resultados que se encuentran siempre dentro del rango $(0, M)$, independientemente de los valores del filtro que se aplique. Esto tiene unas connotaciones muy importantes, al conocerse de antemano que la imagen resultante tras aplicar *Convolución-LIP* contendrá valores dentro de los rangos de visibilidad definidos. Así pues, no será necesario realizar comprobaciones de rangos de cada píxel, ni normalizaciones o reajustes de rangos, muy habituales al aplicar muchos algoritmos en otros modelos. Todo ello significará una reducción del tiempo de cómputo necesario, además de todas las ventajas que ofrece el modelo LIP para la aplicación de algoritmos de procesamiento de imágenes y vídeo.

CRITERIOS PARA LA EVALUACIÓN DE LA CALIDAD DE CONTORNOS

ÍNDICE

7.1	Objetivo principal y Criterios generales	118
7.1.1	Criterios generales	118
7.2	Criterios específicos	118
7.2.1	DOMINIOS del cuestionario	119
7.2.2	Contexto de las tareas de visión	120
7.2.3	Selección de las técnicas de detección de bordes	120
7.2.4	Selección del ámbito de actuación	122
7.2.5	Selección de los parámetros visuales	123
7.3	Imágenes para el cuestionario de evaluación en BAJA VISIÓN	124
7.4	Conclusiones	134
7.4.1	Contexto del capítulo	134

Todas las técnicas nuevas que se proponen en la literatura científica deben presentar una evaluación de sus resultados, en la que el método propuesto sea comparado con otros ya consolidados. Esta evaluación termina reduciéndose a uno o varios valores numéricos, que sirven como resumen de la calidad general de dicho método, y que se encuentran normalizados para poder compararlos con los de otras alternativas. En definitiva, el objetivo es determinar qué técnica, método o algoritmo es el *mejor*.

Como se ha indicado en capítulos anteriores (ver Cap. 3 y 4), las técnicas de procesamiento de imágenes son susceptibles de ser evaluadas de manera subjetiva mediante el uso de cuestionarios de opinión. Además, en muchos casos, para evaluar la “facilidad perceptual” de un método, los cuestionarios y la evaluación subjetiva es la única opción. A lo largo de este capítulo se tratarán diferentes aspectos de la evaluación de las imágenes procesadas, mostrando la dificultad que presenta el desarrollo de los cuestionarios para su evaluación subjetiva. Se seleccionarán los diferentes DOMINIOS de las imágenes utilizados en el cuestionario y se elegirán las diferentes imágenes representativas de cada uno de los mismos. También se especificarán los métodos que se evaluarán mediante cuestionarios.

7.1 OBJETIVO PRINCIPAL Y CRITERIOS GENERALES

En el Capítulo 4 se muestran dos listas de tareas que hay que seguir. En ellas, tanto Burgess [16] como Oppenheim [76], indican como primer paso el de definir el propósito de la investigación de opinión. En la Tab. 7.1 se muestra el objetivo principal de la investigación de la presente Tesis Doctoral.

Tabla 7.1
Objetivo principal del estudio de investigación.

Determinar si los métodos de extracción de bordes basados en el paradigma LIP generan mapas de contornos de mayor calidad para aplicaciones de soporte visual humano que el resto de los métodos de extracción de bordes.

Como se puede observar, el propósito del estudio de investigación presenta términos muy genéricos, que se irán acotando a continuación.

7.1.1 Criterios generales

Tras fijar el objetivo del estudio de investigación, para que la evaluación de los algoritmos de extracción de bordes sea eficaz, hay que fijar *a priori* unos criterios. En el Capítulo 3 se mostró cómo Heath [47] propuso unos criterios para el proceso de evaluación que realizó, que son adaptables a cualquier otro proceso de evaluación de algoritmos de procesamiento de imágenes. Dichos criterios, expuestos en la Tab. 3.3 de la página 60, son los criterios generales que se han seguido en esta Tesis para diseñar los cuestionarios para la evaluación de algoritmos de extracción de bordes. En este trabajo también se han seguido esos criterios con carácter general, puesto que están avalados por un proyecto de Máster en una Universidad de prestigio [47], y por la publicación de varios artículos científicos en congresos internacionales [48] y en revistas de alto impacto [49, 50]. A continuación, se puntualizan todos aquellos aspectos específicos que han sido de interés para el desarrollo del proceso de evaluación.

7.2 CRITERIOS ESPECÍFICOS

La primera precisión sobre la que hay que incidir es la determinación del contexto en el que se trabaja, es decir, especificar la *tarea de visión* en la que se miden las *variables* estadísticas que permiten alcanzar el OBJETIVO FUNCIONAL descrito en la Tab. 7.1. Según Heath [47], “*aunque se podría medir la calidad de los bordes fuera del contexto de una tarea de visión, los resultados obtenidos serían de difícil interpretación en términos de procesamiento de imágenes*”. Por ello, para poder concretar la *tarea de visión*, primero hay que estudiar qué DOMINIOS se consideran dentro del cuestionario. Tras ello, se podrán seleccionar los métodos de extracción de bordes que se evaluarán. Esto permitirá determinar el ámbito de actuación para dicha evaluación. Para ello, será necesario establecer las poblaciones objetivo y ajustar los parámetros visuales para que todos

los métodos se encuentren en situaciones de partida similares para su evaluación.

7.2.1 DOMINIOS del cuestionario

Ampliando los criterios propuestos por Heath [47], éste afirma que “*un sistema de visión específico se debe evaluar tomando un DOMINIO particular y dicha evaluación debe realizarse utilizando imágenes reales representativas del DOMINIO particular en el que se pretende evaluar.*” Para poder seleccionar las imágenes hay que determinar primero los DOMINIOS de los cuestionarios de evaluación en el contexto de esta Tesis Doctoral.

Gran parte de esta tarea se ha realizado en el Capítulo ??, en el que se preseleccionaron 7 DOMINIOS. Esta preselección se ha realizado tomando en consideración únicamente aquellos DOMINIOS que son relevantes dentro de los cuestionarios de evaluación de la BAJA VISIÓN, y que además tienen utilidad dentro del ámbito de trabajo de los métodos propuestos en esta Tesis.

De estas 7 DIMENSIONES preseleccionadas, D₄ (*Sensibilidad y Adaptación a la luz y problemas de iluminación, deslumbramientos y Visión y Conducción nocturna*) es claramente un DOMINIO transversal, ya que no es posible obtener imágenes exclusivamente de este tipo sin que estén cruzadas con el resto de DIMENSIONES.

Se descarta el DOMINIO D₂ (*Agudeza visual en la Visión cercana, lectura de libros, periódicos, listín telefónico, escritura de cheques*), porque los requisitos y parámetros que deben tener los métodos de extracción de contornos para este ámbito difieren mucho de las otras. Esto sugiere que este DOMINIO representa una problemática independiente del resto, que debería ser enfocada de manera explícita y no conjuntamente. Adicionalmente, debido a las características propias de la tarea descrita (*lectura y visión cercana*), ésta requiere una buena iluminación sin sombras, aspecto este último que choca directamente con el DOMINIO transversal D₄.

Continuando con el análisis, la DIMENSIÓN D₁₀ (*Actividad principal: trabajo y tareas del hogar*) puede unificarse con la DIMENSIÓN D₈ (*Cuidado personal: aseo, baño, comer*) en un nuevo DOMINIO, que puede nombrarse como *Visión cercana: Lectura y tareas básicas personales*. Por tanto, de los 7 DOMINIOS preseleccionados en el Capítulo ?? se ha logrado reducir la lista a tan sólo cinco DOMINIOS, que se muestran en la Tab. 7.2. De ellos, los cuatro primeros permiten la obtención de imágenes específicas para cada uno de ellos, mientras que el quinto es transversal a todos los demás. Con esta reducción, se ha conseguido además que todos los cuestionarios analizados en el Cap. ?? se encuentren representados en al menos uno de los DOMINIOS de la lista. En esta tabla también se incluye el nombre abreviado por el que se denominará, por comodidad, cada DOMINIO.

Tabla 7.2
DOMINIOS seleccionados para el cuestionario de BAJA VISIÓN.

1. Visión cercana: Lectura y tareas básicas personales (DOMINIO cercanas).
2. Visión media/larga distancia: Orientación y seguridad en exteriores (DOMINIO exteriores).
3. Visión media/larga distancia: Orientación y seguridad en interiores (DOMINIO interiores).
4. Visión media/larga distancia: Ubicación y reconocimiento de personas (DOMINIO personas).
5. Problemas de iluminación, deslumbramientos y adaptación a la luz. (*Transversal*)

7.2.2 Contexto de las tareas de visión

Por su naturaleza, las imágenes contienen una cantidad de información muy elevada, que el sistema visual humano se encarga de procesar desde los niveles más bajos (extracción de bordes, agrupamiento de texturas y colores, etc.) hasta abstracciones y connotaciones conceptuales de muy alto nivel, como pueden ser, asignaciones de percepciones psico-visuales (por ejemplo, saber si los colores de un determinado objeto son válidos) y/o culturales (por ejemplo, asociar conceptos positivos o negativos a símbolos). Esta gran cantidad de información subyacente en las imágenes hace que no todos los usuarios se fijen en los mismos aspectos al observar la misma imagen. Por ello es conveniente seleccionar una tarea de visión común, que permita al encuestador centrar la atención visual en un aspecto concreto y específico, lo que permite a los usuarios concentrarse en la acción encomendada y olvidarse de los otros estímulos visuales presentes en las imágenes. Así pues, en la Tab. 7.3 se describen las tareas típicas de visión artificial [12, 42, 44, 56]. Teniendo en cuenta los DOMINIOS expuestos en la Tab. 7.2 y las tareas de visión descritas en la Tab. 7.3 se puede concluir que el contexto de la tarea de visión de las preguntas del cuestionario debe ser la tarea general de **RECONOCIMIENTO DE OBJETOS**, y de manera específica, dependiendo de cada imagen en particular, la tarea concreta debería ser la de **DETECCIÓN**, **RECONOCIMIENTO** o **IDENTIFICACIÓN**. Estas tres tareas están consideradas como los tres niveles progresivos del **RECONOCIMIENTO DE OBJETOS**. Esta afirmación se ve reforzada con el trabajo de Abdou y Dusassoy [1], en el que los observadores humanos que procesan las imágenes catalogan las tareas de **RECONOCIMIENTO DE OBJETOS** en cuatro categorías: **DETECCIÓN**, **ORIENTACIÓN**, **RECONOCIMIENTO** e **IDENTIFICACIÓN**.

7.2.3 Selección de las técnicas de detección de bordes

Tomando la clasificación expuesta en el Capítulo 2, los métodos para la extracción de bordes se dividen en tres grandes bloques:

Tabla 7.3
Tareas típicas de
visión artificial.

-
- A. **RECONOCIMIENTO DE OBJETOS:** *Determinar si en una escena está presente algún objeto específico, alguna característica o alguna actividad concreta.*
- DETECCIÓN** *Búsqueda dentro de la imagen de alguna condición especial, por ejemplo, invasión de un carril por un vehículo o búsqueda de células infectadas en una muestra.*
- ORIENTACIÓN** *Detección de alguna parte de un objeto que permita determinar, al menos, la orientación de dicho objeto.*
- RECONOCIMIENTO** *Determinación de la posición de un objeto genérico previamente “aprendido”.*
- IDENTIFICACIÓN** *Reconocimiento de una instancia específica de un objeto, por ejemplo, una determinada cara o una huella dactilar.*
- B. **ESTIMACIÓN DEL MOVIMIENTO:** *Mecanismo para estimar la velocidad relativa de los diversos objetos de una escena.*
- SEGUIMIENTO** *Capacidad para rastrear los movimientos de un objeto en una escena (típicamente, un vehículo o una persona).*
- CALIBRACIÓN DE CÁMARA** *Determinación del movimiento rígido 3D de la cámara que está tomando las imágenes.*
- C. **RECONSTRUCCIÓN DE ESCENAS:** *Recreación 3D de una escena a partir de varias imágenes 2D.*
- D. **MEJORA DE LA CALIDAD DE LAS IMÁGENES:**
- RESTAURACIÓN** *Eliminación de ruido, eliminación de emborronamiento, etc.*
- MEJORA** *Técnicas que mejoran la calidad percibida de las imágenes: aumento del contraste, etc.*
-

- Basados en el Gradiente.
- Basados en el Laplaciano.
- Basados en operadores Morfológicos.

De los tres tipos de extractores, el último (*basados en operadores Morfológicos*) no se ha contemplado en esta Tesis Doctoral, puesto que no utiliza el operador *convolución* y, por tanto, se aleja del planteamiento original del trabajo. Para poder realizar una evaluación realista, se ha escogido un algoritmo representativo de cada uno de los bloques, y uno más adicional, el algoritmo de *Canny*, que está considerado como uno de los métodos de más alta calidad para la extracción de contornos. Junto a estos, se han tomado los respectivos algoritmos implementados bajo el paradigma LIP, descritos con mayor profundidad en el Cap. 9.

Tabla 7.4
Técnicas seleccionadas para su evaluación.

- A. Basado en el Gradiente:
 1. SOBEL.
 2. LIP-SOBEL.
- B. Basado en el Laplaciano:
 1. LoG.
 2. LIP-LoG.
- C. Alta precisión:
 1. CANNY.
 2. LIP-CANNY.

7.2.4 Selección del ámbito de actuación

Como se ha expuesto con anterioridad, es necesario fijar el ámbito de actuación para evaluar la calidad de las técnicas propuestas utilizando el paradigma LIP. Con esto se consigue fijar los objetivos que se desean medir en dicha evaluación y que sirven para determinar la estructura y contenidos de la encuesta de opinión que se utilizará para evaluar la calidad de dichas técnicas. En este ámbito, se presentan dos objetivos principales que se desean contrastar mediante la evaluación subjetiva, que se muestran en la Tab. 7.5.

Tabla 7.5
Objetivos principales de la encuesta.

- A. Los métodos desarrollados bajo LIP presentan un mejor comportamiento para la detección de objetos en sombras para usuarios sin problemas particulares de visión.
- B. Los pacientes con BAJA VISIÓN prefieren las imágenes obtenidas con los métodos desarrollados bajo LIP frente a las obtenidas con los métodos tradicionales.

Los estudios de investigación mediante cuestionarios necesitan identificar las poblaciones de individuos a los que se les realizarán las preguntas de los cuestionarios. Como se puede observar del estudio de

los objetivos principales, cada uno de ellos tiene una población de interés diferente, por lo que, se pueden detectar dos grupos poblacionales claramente diferenciados, que se muestran en la Tab. 7.6.

- A. Usuarios con nivel de VISIÓN ESTÁNDAR.
- B. Pacientes con BAJA VISIÓN.

Tabla 7.6
Tipologías de grupos poblacionales de la encuesta.

7.2.5 Selección de los parámetros visuales

Para que las comparaciones entre los diversos algoritmos se puedan realizar en términos de igualdad, todos los algoritmos y métodos deben proporcionar bordes de similar grosor y calidad. Como elemento de normalización, en esta Tesis Doctoral, se ha incorporado a todas las técnicas un mecanismo de supresión de bordes no maximales y de posterior histéresis de contornos. Por último, en todos los métodos se ha decidido aplicar un filtrado morfológico final para obtener imágenes de bordes con características homogéneas. Estas características se muestran en la Tab. 7.7.

- A. Todos los bordes deben tener 1 píxel de grosor.
- B. Deben detectarse únicamente píxeles de bordes no espúreos.

Tabla 7.7
Características homogéneas de todos los métodos evaluados en la encuesta.

Para la detección de bordes es necesario establecer un umbral que determine si el valor procesado de un determinado píxel es considerado borde o no. La selección de dicho umbral es un aspecto delicado, puesto que una mala selección del umbral puede provocar un resultado pésimo en la evaluación. Además, aunque en la literatura científica no se ha encontrado ningún trabajo que indique la cantidad máxima de contornos que podría admitir un usuario para no verse saturado, los pacientes con BAJA VISIÓN informan que es preferible una menor cantidad de bordes antes que una alta precisión de los mismos. De manera heurística, mediante ensayos realizados por un conjunto de expertos en la materia, se ha estimado que el porcentaje máximo de contornos ha de ser, en todo caso, menor a un 20% de los píxeles totales de la imagen. Utilizando este porcentaje, para cada imagen, se determinará un valor del umbral, de tal manera, que el número de píxeles de bordes seleccionados en la imagen sea menor a dicho porcentaje. Así pues, los parámetros se ajustarán teniendo en cuenta las características indicadas en la Tab. 7.8.

Tabla 7.8
Determinación de
umbrales.

- A. El número de píxeles de contornos en el mapa final se debe mantener en valores porcentuales bajos ($\leq 20\%$) en relación con el tamaño de la imagen.
- B. Un experto realizará la selección del valor exacto del umbral.

7.3 IMÁGENES PARA EL CUESTIONARIO DE EVALUACIÓN EN BAJA VISIÓN

La selección de las imágenes utilizadas en el cuestionario se ha realizado en función de los DOMINIOS seleccionados previamente (ver Tab. 7.2). Inicialmente se ha partido de un total de 114 imágenes de distintos DOMINIOS:

- VISIÓN CERCANA: 32 imágenes.
- ORIENTACIÓN EN INTERIORES: 24 imágenes.
- ORIENTACIÓN EN EXTERIORES: 39 imágenes.
- RECONOCIMIENTO DE PERSONAS: 19 imágenes.

Estas imágenes provienen de bancos de imágenes públicas, o bien, han sido tomadas por el propio autor de esta Tesis. Cada una de estas imágenes presenta un elemento principal que permite catalogarla dentro de uno de los DOMINIOS. Además, todas las imágenes seleccionadas contienen *iluminación* no uniforme, sombras u otras características propias del DOMINIO transversal expuesto en la Tab.7.2.

Utilizando esta primera selección de imágenes, se han aplicado las diferentes técnicas seleccionadas previamente (*Sobel*, *LIP-Sobel*, *LoG*, *LIP-LoG*, *Canny*, *LIP-Canny*) con las restricciones de los parámetros previamente indicadas (*4 niveles de porcentaje: 5%, 10%, 15% y 20% de contornos en las imágenes*), generándose un conjunto de 2736 imágenes de bordes.

Todas estas imágenes de bordes han sido analizadas por expertos en procesamiento de imágenes, para descartar todas aquellas que no proporcionan resultados aceptables (por exceso de ruido o por insuficiencia de contornos), reduciendo el conjunto de imágenes originales a 20, con un límite de 5 imágenes por DOMINIO. Las Figuras 7.1 a 7.20 muestran el conjunto reducido de imágenes originales utilizado para generar las imágenes de bordes del cuestionario de evaluación. Todas estas imágenes están disponibles en: <http://www.uco.es/users/el2pamuj/research/LowVision/>.



Figura 7.1
Imagen 1 original del
DOMINIO personas.



Figura 7.2
Imagen 2 original del
DOMINIO personas.



Figura 7.3
Imagen 3 original del
DOMINIO personas.

Figura 7.4
Imagen 4 original del
DOMINIO personas.



Figura 7.5
Imagen 5 original del
DOMINIO personas.





Figura 7.6
Imagen 1 original del
DOMINIO cercanas.



Figura 7.7
Imagen 2 original del
DOMINIO cercanas.

Figura 7.8
Imagen 3 original del
DOMINIO cercanas.



Figura 7.9
Imagen 4 original del
DOMINIO cercanas.





Figura 7.10
Imagen 5 original del
DOMINIO cercanas.



Figura 7.11
Imagen 1 original del
DOMINIO interiores.

Figura 7.12
Imagen 2 original del
DOMINIO interiores.



Figura 7.13
Imagen 3 original del
DOMINIO interiores.





Figura 7.14
Imagen 4 original del
DOMINIO interiores.



Figura 7.15
Imagen 5 original del
DOMINIO interiores.

Figura 7.16
Imagen 1 original del
DOMINIO exteriores.



Figura 7.17
Imagen 2 original del
DOMINIO exteriores.



Figura 7.18
Imagen 3 original del
DOMINIO exteriores.





Figura 7.19
Imagen 4 original del
DOMINIO exteriores.



Figura 7.20
Imagen 5 original del
DOMINIO exteriores.

7.4 CONCLUSIONES

En este capítulo se han mostrado tanto los objetivos generales como los específicos utilizados para el diseño de los experimentos y de los cuestionarios para evaluar los resultados. El objetivo principal del estudio de investigación es determinar *si los métodos de extracción de bordes basados en el paradigma LIP generan mapas de contornos de mayor calidad para aplicaciones de soporte visual humano que el resto de los métodos de extracción de bordes*. Para lo cual, se obtendrán imágenes de bordes, que serán evaluadas por humanos. Este proceso lleva a tomar algunas decisiones de diseño más específicas. Primero, se han estudiado los cuestionarios de evaluación de BAJA VISIÓN y mediante un exhaustivo proceso de análisis y síntesis, se han seleccionado 4 DOMINIOS específicos, entre los que se escogerán las imágenes para los experimentos, y un DOMINIO transversal, que afecta al resto de DOMINIOS. Los DOMINIOS seleccionados se muestran en la Tab. 7.2. Para evaluar la calidad de las imágenes de bordes, los encuestados deben realizar una cierta *tarea de visión* utilizando como entrada imágenes procesadas, pertenecientes a los distintos DOMINIOS. Sobre cada imagen de bordes, se solicita a los encuestados que realice una determinada *tarea de visión*:

- DETECCIÓN de objetos.
- RECONOCIMIENTO de objetos.
- IDENTIFICACIÓN de objetos.

En la Tab. 7.4 se muestran los algoritmos de extracción de bordes que se han seleccionado para la evaluación mediante cuestionarios.

También se han especificado los parámetros visuales comunes a todos los métodos que se evaluarán. Estos parámetros permitirán comparar, con características normalizadas y homogéneas, los resultados de todos los algoritmos de extracción de bordes, sin favorecer ni perjudicar ninguno de los métodos evaluados. Los parámetros comunes seleccionados se muestran en las Tab. 7.7 y 7.8.

La evaluación de la calidad de los contornos generados por estos algoritmos bajo estos parámetros se realiza en dos ámbitos, mostrados en la Tab. 7.5. Estos dos ámbitos representan los objetivos funcionales primarios de la encuesta.

La cumplimentación de estos objetivos se realizará por 2 grupos poblacionales diferenciados:

- Usuarios con nivel de VISIÓN ESTÁNDAR.
- Pacientes con BAJA VISIÓN.

Para finalizar, se han mostrado las 20 imágenes seleccionadas, con 5 imágenes de cada uno de los cuatro DOMINIOS.

7.4.1 Contexto del capítulo

Tras exponer en el capítulo anterior el nuevo operador *Convolución-LIP*, en éste se han expuesto los criterios generales y específicos que

permiten el desarrollo de los algoritmos de extracción de bordes bajo el modelo LIP, con los que se realizarán los experimentos y la evaluación de los mismos.

Como se ha comentado a lo largo de la Tesis Doctoral, la evaluación se realiza mediante cuestionarios. Sin embargo, al no existir una metodología para el diseño de cuestionarios sobre calidad de imágenes procesadas, la construcción de este tipo de encuestas no es trivial. En capítulos posteriores, se proponen un conjunto de recomendaciones generales para el diseño de cuestionarios que involucren contenidos multimedia, gracias al cual se podrán construir encuestas que permitan catalogar, según su calidad visual, los métodos de extracción de bordes.

RECOMENDACIONES PARA EL DISEÑO DE CUESTIONARIOS DE CALIDAD DE CONTORNOS

ÍNDICE

8.1	Sesgo por finalidad	138
8.2	Recomendaciones generales sobre las preguntas	139
8.2.1	Estructura de las preguntas	139
8.2.2	Imágenes respuesta	140
8.2.3	Ponderación de preguntas	140
8.3	Tipos de preguntas con elementos visuales	141
8.3.1	Preguntas de SELECCIÓN O DICOTÓMICAS	141
8.3.2	Preguntas de PUNTUACIÓN	142
8.3.3	Preguntas de ORDENAMIENTO	143
8.3.4	Preguntas de MAPEADO ESPACIAL Y TEMPORAL	143
8.4	Recomendaciones sobre la estructura general del cuestionario	145
8.4.1	Reutilización de imágenes	145
8.4.2	Uso de imágenes reales	146
8.4.3	Número de <i>imágenes respuesta</i> por pregunta	147
8.5	Sistema Web para cuestionarios multimedia	148
8.5.1	Preguntas de PUNTUACIÓN	150
8.5.2	Preguntas de SELECCIÓN SIN REPETICIÓN	150
8.5.3	Preguntas de SELECCIÓN CON REPETICIÓN	151
8.5.4	Preguntas de MAPEADO ESPACIAL Y TEMPORAL	151
8.6	Conclusiones	152
8.6.1	Conclusiones de la Parte II	152

Muchos autores indican que no es posible proporcionar un único e infalible conjunto de reglas, protocolos o metodologías para la correcta construcción o diseño de cuestionarios, que eliminen completamente todo tipo de sesgo. En general, lo que se proponen son recomendaciones en el diseño de las preguntas y de los cuestionarios para minimizar dicho sesgo. Como se ha indicado en capítulos anteriores, existen algunas recomendaciones de diseño en el plano lingüístico para la realización de cuestionarios. Sin embargo, no existen recomendaciones para cuestionarios en los que se incluyan elementos multimedia. Se inicia identificando el nuevo *sesgo por finalidad* y describiendo algunas recomendaciones para mitigar su efecto. Se continúa aportando algunas recomendaciones de diseño para cuestionarios que involucren preguntas con elementos visuales, y en particular, para la evaluación subjetiva de la calidad de algoritmos de extracción de contornos. Para finalizar este capítulo, se presenta una herramienta Web para la creación y difusión de los cuestionarios, que permite contestarlos on-line, y que facilita al investigador la recolección de las respuestas.

8.1 SESGO POR FINALIDAD

El *Sesgo por finalidad* se produce cuando los encuestados conocen los objetivos asociados a la pregunta que están respondiendo y, por tanto, las respuestas pueden estar influenciadas por ello. En algunos casos, el efecto producido por el *Sesgo de deseabilidad social* es similar, puesto que los encuestados no contestan según sus opiniones propias, sino las que los encuestados creen que son las que mayor aprobación social tendrán. En cierto modo, los encuestados escogen las opciones influenciados por lo que ellos creen que el entrevistador desea que ellos contesten. En estos casos, las recomendaciones de diseño indican que las preguntas que puedan sufrir de algún tipo de *Sesgo de deseabilidad social* deben hacerse en un tono neutral y mediante preguntas indirectas, de tal forma que sea más complicado percibir la intencionalidad (los objetivos subyacentes asociados) de dichas preguntas.

Las encuestas que tratan de estimar la calidad de algún proceso suelen depender de la finalidad con la que se aplique dicho proceso. Es decir, que un mismo proceso puede proporcionar un resultado muy positivo para una aplicación determinada, y muy negativo para otras. Por ejemplo, un medicamento puede ser muy eficaz para tratar una enfermedad y, a la vez, estar contraindicado para otras. Este sesgo puede aparecer en diferentes escenarios: cuando el entrevistado conoce (directamente informado por el entrevistado o porque ha sido capaz de deducirlo) la finalidad del proceso sobre el que se le está encuestando, o cuando cree conocer dicha finalidad y realmente se confunde a sí mismo, proporcionando respuestas totalmente erróneas. Normalmente, suele estar causado por una incorrecta redacción de las preguntas por parte del entrevistador, no dejando claramente especificada la finalidad de cada pregunta.

En el caso de las encuestas de calidad, en muchos casos hay que guiar al encuestado sobre qué debe observar o analizar, proporcionándole en estos casos el objetivo de la pregunta. Por ejemplo, si en una cata ciega de un nuevo producto se desea que los usuarios den su opinión sobre un determinado aspecto del mismo, habría que indicarles que se deben centrar en un aspecto en particular, ya que no se desea conocer otros aspectos de ese producto. Concretando el ejemplo, si el producto fuese una bebida, se podría preguntar en exclusiva sobre el dulzor, la acidez, la textura, etc. Sin embargo, puede que ese aspecto en particular no sea el que más llame la atención a los encuestados, o incluso, que si no se le hubiese preguntado específicamente por él, habría pasado desapercibido. En estos casos, aunque los encuestados proporcionen las respuestas basadas en su propia opinión, éstas están influidas por el *Sesgo por finalidad*.

En la mayoría de los casos en los que aparece este tipo de sesgo, éste se introduce debido a que el diseñador de la encuesta necesita indicar a los encuestados el objetivo que se está buscando en cada pregunta, para centrar la atención de los encuestados en dicho aspecto, descartando otros elementos presentes en dicha pregunta. Se plantean dos posibilidades:

- A. No informar a los encuestados de la finalidad de cada pregunta, permitiendo que las respuestas de los encuestados no se ajusten a los objetivos de cada pregunta.
- B. Informar a los encuestados de la finalidad en cada pregunta, asumiendo que se está introduciendo este tipo de sesgo en todas las preguntas.

Con la primera posibilidad, se elimina el *Sesgo por finalidad*, ya que se está omitiendo la información de la finalidad de las preguntas, por lo que no se influye a los encuestados. Sin embargo, no hay garantía de que los encuestados proporcionen respuestas que sirvan para contestar a los objetivos asociados a cada pregunta.

Por el contrario, con la segunda posibilidad se introduce explícitamente este sesgo en todas las preguntas. De esta manera se reduce la libertad de los encuestados en las posibles respuestas, pero se intenta garantizar un nivel de conocimiento base que permita a los encuestados proporcionar respuestas que sean interesantes para el evaluador en el contexto para el que ha desarrollado la pregunta. Esta segunda opción es admisible si se informa a todos los encuestados de los objetivos de cada una de las preguntas del cuestionario. No es conveniente mezclar en un mismo cuestionario preguntas en las que se indique el objetivo de las mismas con otras preguntas en las que esto no se indique, ya que, al estar más guiadas unas respuestas que otras, los resultados no son comparables.

8.2 RECOMENDACIONES GENERALES SOBRE LAS PREGUNTAS

Con los cuestionarios indicados en esta Tesis se intenta evaluar de manera subjetiva qué algoritmo de procesamiento de imágenes se comporta de mejor manera ante un determinado entorno de funcionamiento (que vienen determinado por los DOMINIOS del cuestionario). A continuación, se muestra la estructura que se recomienda que tenga un cuestionario.

8.2.1 Estructura de las preguntas

Se considera que un cuestionario está compuesto por una o varias preguntas; y cada pregunta está compuesta por:

- A. Un texto de introducción y de pregunta.
- B. Una imagen original [Optativa].
- C. Una o varias *imágenes respuesta*, cada una con su zona de respuesta asociada.

Se ha considerado que lo más eficiente para este tipo de cuestionarios es no mezclar preguntas que contengan elementos visuales con preguntas de tipo exclusivamente lingüístico. Si se mezclan ambos tipos de preguntas, la evaluación se complica mucho para el entrevistador, pues, como se verá más adelante, cada pregunta tiene ponderada su

contribución al resultado final, y conseguir dicho factor de ponderación en preguntas exclusivamente lingüísticas es muy complejo. Además, se han realizado diversas pruebas con cuestionarios experimentales (pruebas de pilotaje inicial), que han mostrado que los entrevistados mostraban mayores dificultades al responder cuestionarios con mezcla de ambos tipos de preguntas (lingüísticas y visuales) que los cuestionarios con preguntas de un único tipo. Los entrevistados argumentaron que les “costaba” ir cambiando de formato de pregunta cuando ya se habían acostumbrado a responder preguntas de un determinado tipo.

8.2.2 *Imágenes respuesta*

Las *imágenes respuesta* son las distintas imágenes de respuesta asociadas a cada pregunta, donde se busca que el entrevistado responda a la cuestión planteada en el “texto de introducción y de pregunta”. En dicho texto se solicita al entrevistado que realice alguna tarea de visión (DETECCIÓN, ORIENTACIÓN, RECONOCIMIENTO O IDENTIFICACIÓN) utilizando en cada caso la *imagen respuesta* correspondiente. Cada una de ellas tendrá asociado un único algoritmo de procesamiento de imágenes.

8.2.3 *Ponderación de preguntas*

No todas las preguntas tienen el mismo valor para obtener el resultado final, puesto que hay algunas preguntas que pueden ser discriminatorias, mientras que otras pueden matizar el grado de especialización. Tómese como ejemplo, una pregunta en la que se solicite que se indiquen aquellas *imágenes respuesta* en las que se pueda detectar un determinado objeto, sin importar si se detecta mejor o peor. Todas aquellas *imágenes respuesta* en las que no sea apreciable dicho objeto obtendrán baja puntuación, y por lo tanto, los algoritmos de procesamiento de imágenes asociados a esas *imágenes respuesta* conseguirán pocos puntos. Pero además, desde un punto de vista global, esa pregunta es altamente discriminatoria, es decir, que para la puntuación global de los algoritmos de procesamiento de imágenes, existe una gran diferencia en el grado de colaboración en la puntuación final de los casos en los que los entrevistados pueden detectar el objeto frente a los casos en los que no pueden detectarlo. Por tanto, cada pregunta colaborará con el resultado final con un cierto grado, especificado utilizando valores de ponderación, que podrán variar según las distintas clases de respuestas posibles.

A pesar de que este comportamiento es altamente deseable, una ponderación de todas las preguntas de un cuestionario que sea satisfactoria y que no presente ni *sesgo* ni favorezca *a priori* ningún método es muy difícil de conseguir. Para poder simplificar la tarea del diseño de la encuesta, en esta Tesis Doctoral no se aplicarán distintas ponderaciones, sino que todas las preguntas colaborarán en la misma proporción en el cálculo del resultado final.

8.3 TIPOS DE PREGUNTAS CON ELEMENTOS VISUALES

Los posibles tipos de preguntas que se consideran en los cuestionarios de calidad de bordes en imágenes son los de:

- A. Preguntas con respuestas de SELECCIÓN (o DICOTÓMICAS).
 1. CON REPETICIÓN (Varias *imágenes respuesta* válidas).
 2. SIN REPETICIÓN (Sólo una *imagen respuesta* válida).
- B. Preguntas con respuestas de PUNTUACIÓN.
 1. CON REPETICIÓN (Varias *imágenes respuesta* pueden tener la misma puntuación).
 2. SIN REPETICIÓN (Cada *imagen respuesta* tiene una puntuación, y dicha puntuación no ha sido asignada a ninguna otra *imagen respuesta*).
- C. Preguntas con respuestas de ORDENAMIENTO.
 1. CON REPETICIÓN (Varias *imágenes respuesta* pueden tener el mismo valor de asignación).
 2. SIN REPETICIÓN (Cada *imagen respuesta* tiene un valor único de ordenamiento).
- D. Preguntas con respuesta de MAPEADO ESPACIAL Y TEMPORAL (o MAPPING).

8.3.1 Preguntas de SELECCIÓN o DICOTÓMICAS

Mediante este tipo de preguntas se busca encontrar en cuáles de las *imágenes respuesta* el entrevistado puede realizar la tarea de visión encomendada en el texto de la pregunta. Para desambiguar cualquier posible duda que pudiera tener el entrevistado, el entrevistador debe especificar con la mayor claridad posible en el texto de la pregunta en qué casos puede el entrevistado seleccionar la *imagen respuesta* en función de la realización de la tarea de visión (por ejemplo, si puede realizar perfectamente la tarea de visión encargada sin ninguna dificultad, o en otro caso distinto, si puede realizar al menos mínimamente la tarea de visión aunque con un cierto grado de dificultad, etc.).

SIN REPETICIÓN Utilizando el mecanismo SIN REPETICIÓN, el entrevistador consigue un beneficio directo: obliga al entrevistado a escoger sólo una de las *imágenes respuesta*, a pesar de que quizás éste pudiera realizar la tarea de visión encomendada en varias de ellas. El principal handicap que comentan los entrevistados sobre esta opción es que es poco natural, debido a que les fuerza a escoger, aún cuando no lo desean. Un beneficio adicional que se consigue con este mecanismo es que éste se puede considerar un selector elitista, es decir, que permite discriminar con mayor claridad cuál es considerado el “mejor” algoritmo (asociado a una *imagen resultado*), ya que éste obtiene toda la puntuación, mientras que el resto no obtiene ningún punto.

CON REPETICIÓN Por el contrario, el mecanismo CON REPETICIÓN permite un mayor grado de libertad al entrevistado, ya que éste no tiene que escoger sólo una *imagen respuesta*. Esta opción es mucho más natural, ya que si no existen excesivas diferencias entre los resultados de los distintos algoritmos que se están evaluando, es mucho más realista para un entrevistado afirmar que es capaz de realizar una determinada tarea de visión con varias *imágenes respuesta* similares. Sin embargo, al contrario que en el mecanismo anterior, éste permite una menor discriminación al obtener igual puntuación varios de los algoritmos (pudiendo incluso seleccionar todas las *imágenes respuesta*).

8.3.2 Preguntas de PUNTUACIÓN

Gracias a estas preguntas, el entrevistador puede conocer el grado de calidad intrínseca percibida por los entrevistados asociada a la utilización de cada *imagen respuesta* para realizar una tarea de visión (especificada en el “texto de la pregunta”).

La utilización de varias imágenes permite a los entrevistados compararlas, por lo que en esos casos, la puntuación que se asigne se puede considerar más como el grado de comparación entre las imágenes que como la evaluación individual de cada una. Al permitirse la comparación entre las imágenes, los entrevistados al puntuar cada una de ellas, incorporan más información que la incluida en la propia imagen, ya que pueden mirar el resto de imágenes. En muchos casos, la puntuación de cada *imagen respuesta* se realiza valorando cada imagen con una imagen “virtual” que cada entrevistado se construye de manera heurística en su propia mente utilizando información visual de cada *imagen respuesta*. Por tanto, en las preguntas que se utilicen más de una *imagen respuesta* el entrevistador debe admitir la existencia de un cierto grado de *sesgo por finalidad*, que debe ser minimizado mediante la inclusión de otras preguntas en las que no se muestren múltiples *imágenes respuesta*.

SIN REPETICIÓN Utilizando el mecanismo SIN REPETICIÓN, el entrevistador fuerza a los entrevistados a asignar diferente puntuación a todas las *imágenes respuesta*. Si la puntuación es un valor real, esta opción no tiene utilidad práctica, por la facilidad que existe para modificar la puntuación decimal. Por el contrario, si es un valor entero, se provoca una cierta distanciaci3n entre las *imágenes respuesta*, lo que permite una mayor discriminaci3n en las respuestas. Sin embargo, esa diferenciaci3n puede ser totalmente ficticia, en el sentido que los entrevistados pueden encontrar una diferencia entre las *imágenes respuesta* menor que la que le obliga a la asignaci3n sin repetic3n y con valores enteros.

CON REPETICIÓN Por el contrario, el mecanismo CON REPETICIÓN permite una mayor flexibilidad a los entrevistados para asignar puntuaci3n. Al permitir la repetic3n, a veces, los entrevistados tienden a agrupar puntuaciones, es decir, asignan la misma puntuaci3n a todas las *imágenes respuesta* que tienen una calidad elevada, sin precisar si en alguna de ellas la calidad es mayor o menor.

8.3.3 Preguntas de ORDENAMIENTO

Gracias a estas preguntas, el entrevistador puede conocer el orden de calidad percibida comparada por los entrevistados asociada a la utilización de cada *imagen respuesta* para realizar una tarea de visión (especificada en el “texto de la pregunta”). Como su propio nombre indica, el entrevistado “ordena” de mejor a peor las imágenes, asignando un valor numérico a cada *imagen respuesta* que permite organizarlas comparativamente. Con este tipo de preguntas, el entrevistador no tiene garantía de que el entrevistado haya sido capaz de realizar la tarea de visión solicitada: sólo se solicita que ordene de mejor a peor las *imágenes respuesta*, cuando quizás el entrevistado no haya sido capaz de realizar la tarea encomendada, aunque de manera subjetiva es capaz de indicar con cuál se siente más cómodo.

Se puede afirmar que este tipo de preguntas es menos exigente para el entrevistado que las de PUNTUACIÓN, ya que al no tener que puntuar, no tiene que comparar mentalmente cada *imagen respuesta* con la *imagen respuesta* “perfecta” (que, previamente, el entrevistado se habrá formado en su mente); le basta con ordenar las existentes. Debido a la simplificación en la evaluación, este tipo de preguntas no proporcionan resultados excesivamente discriminantes. Un cuestionario realizado exclusivamente con este tipo de preguntas sólo proporcionaría *tendencias* generales y para obtener resultados más concretos, estas preguntas se deben formular añadidas a otras de otros tipos.

SIN REPETICIÓN Utilizando el mecanismo SIN REPETICIÓN, el entrevistador obliga que el entrevistado no pueda indicar que dos o más *imágenes respuesta* son iguales o tienen la misma calidad para realizar la “tarea de visión”. Este aspecto incrementa ligeramente la discriminación, pero los entrevistados suelen mostrarse molestos por tener que escoger entre uno en particular y se pueden producir pequeños *sesgos* con respecto a la posición de las *imágenes respuesta* entre las que tiene que escoger (existiendo una leve preferencia hacia las que se encuentran a la izquierda, probablemente por ser la tendencia natural de la escritura occidental, de izquierda a derecha).

CON REPETICIÓN Por el contrario, el mecanismo CON REPETICIÓN permite mayor libertad a los entrevistados para clasificar las *imágenes respuesta*, a costa de que dichas respuestas suelen tener menor discriminación que el tipo anterior.

8.3.4 Preguntas de MAPEADO ESPACIAL Y TEMPORAL

Mediante este tipo de preguntas, el entrevistador puede comprobar la precisión espacial y el tiempo que tardan los entrevistados en realizar las “tareas de visión” encomendadas. Estas preguntas, también llamadas de MAPPING, sólo admiten una *imagen respuesta*, sobre la cual el entrevistado debe marcar en la posición en la que el propio entrevistado estima que se encuentra la característica solicitada en el texto de la pregunta. El entrevistador deberá tener señalado en una copia propia de la *imagen respuesta*, que no mostraría en ningún caso al entrevistado,

la posición exacta de la característica buscada, a lo que se ha dado en denominar *punto* (o *zona*) *objetivo*. El encuestador podrá señalar zonas de precisión, es decir, zonas alrededor del *punto objetivo* que podrán tener distinta ponderación. Mediante esta opción se controla la precisión espacial. En el caso de que la característica buscada sea marcable de manera puntual, bastará con que el entrevistado señale dicho punto (por ejemplo, con el puntero del ratón). Si la característica fuese una zona, el entrevistado tendría que marcar varios puntos que delimitasen dicha zona, para lo cual el sistema (si fuese por computador) podría facilitar el trabajo del entrevistado permitiendo el manejo automático de estructuras poligonales. Esta opción, sin embargo, presenta un inconveniente que es el manejo de los errores (áreas que se encuentren fuera de la zona marcada por el entrevistador como *zona objetivo*).

En muchas aplicaciones, la precisión espacial es importante, pero dentro de unos límites temporales. Es decir, existen situaciones en las que el entrevistado debería responder antes de un cierto tiempo, a partir del cual, la consideración de la respuesta podría estar muy degradada o incluso puede no ser aceptable en absoluto (a pesar de que la misma fuese perfecta en cuanto a la precisión espacial). Para controlar esta opción el sistema debe controlar el tiempo de respuesta de los entrevistados ante la presentación del estímulo (la *imagen respuesta*). En estos casos, se debe mostrar el texto de la pregunta a los entrevistados manteniendo la *imagen respuesta* oculta hasta que éste indique que está preparado y se empieza a contar el tiempo. El encuestador debe tener una función matemática que permita ponderar el tiempo de respuesta del entrevistado.

Utilizando este tipo de preguntas, el encuestador puede cruzar datos con otras preguntas y comprobar si realmente los entrevistados están realizando la "tarea de visión" encomendada. Esto permite controlar incorrecciones por parte de los encuestados. Por ejemplo, en otros tipos de preguntas, los entrevistados podrían afirmar que son capaces de detectar un cierto objeto en una determinada *imagen respuesta*, y sin embargo, al proponerles que señalen donde se encuentra dicho objeto mediante una pregunta de MAPPING, el entrevistado señala en otra zona distinta o no con demasiada precisión, o tarda mucho tiempo en determinar la posición exacta. Esto significaría que la respuesta que dieron a la otra pregunta no era demasiado veraz, y por tanto, habría que tener mucha cautela al incluir dichas respuestas en el cálculo final. El principal inconveniente que presenta este tipo de preguntas es que sólo se pueden evaluar una *imagen respuesta* cada vez. También hay que indicar que son un poco más difíciles y algo más tediosas de contestar que el resto.

Para finalizar, se ha de indicar que este tipo de preguntas se presenta como una aportación especial de esta Tesis, y se ha diseñado haciendo uso específico del soporte informático, aunque con algunas pequeñas adaptaciones podría utilizarse en encuestas "en papel".

8.4 RECOMENDACIONES SOBRE LA ESTRUCTURA GENERAL DEL CUESTIONARIO

A continuación, se analizan algunos aspectos que tienen una influencia en el diseño del cuestionario y se propondrán una serie de recomendaciones sobre la estructura del cuestionario. La mayoría de estas recomendaciones intentan reducir el *sesgo* que el uso de contenidos visuales puede introducir en el cuestionario en general y en cada pregunta en particular.

8.4.1 Reutilización de imágenes

Uno de los primeros aspectos que se consideran cuando se diseña un cuestionario con contenidos visuales, es la reutilización de esos mismos contenidos visuales. Esta reutilización se presenta en dos vertientes, utilizar exactamente las mismas *imágenes respuesta*, obtenidas utilizando el mismo método de procesamiento de imágenes con los mismos parámetros, o bien, utilizar la misma imagen original y obtener diferentes *imágenes respuesta* (ya sea variando el método de procesamiento de imágenes, modificando los parámetros aplicados o ambos casos) que son representaciones visuales de la misma imagen original.

Se ha podido observar que los entrevistados *recuerdan* las características de las imágenes reutilizadas, produciéndose un *sesgo* en la respuesta de los mismos. Para comprobar este comportamiento se realizó un experimento en los que los entrevistados tenían que puntuar la calidad de los mapas de bordes de ciertas *imágenes respuesta*. El cuestionario estaba compuesto por 5 preguntas. En cada pregunta se aplicaba un mismo método de procesamiento de bordes sobre la misma imagen original, variando los parámetros de dicho método, obteniendo las *imágenes respuesta* que se comparaban en cada pregunta. Sin conocimiento de los entrevistados, la primera pregunta y la última pregunta del cuestionario eran idénticas y mostraban exactamente las mismas *imágenes respuesta*, incluyendo el mismo orden de presentación de las mismas. Se pudo observar que todos los entrevistados respondieron de manera análoga en la primera y en la última pregunta, sin embargo, en general tardaron mucho menos en contestar la última pregunta que la primera, a pesar de que eran iguales. Tras finalizar la encuesta, se les preguntó porqué habían tardado menos en responder a la última pregunta que a la primera y todos ellos respondieron que “*ya conocían las imágenes*”.

En el caso de uso de preguntas con respuestas de MAPEADO ESPACIAL Y TEMPORAL, dónde tanto la precisión de la respuesta como el tiempo de respuesta influyen en la calidad de la misma, la reutilización de imágenes es un elemento crítico. Si utilizando este tipo de preguntas se pretende evaluar la calidad de varios métodos o del mismo método con diferentes parámetros, es imprescindible utilizar la misma imagen original. Para estudiar la influencia de la reutilización de la misma imagen original en preguntas con respuestas de MAPEADO ESPACIAL Y TEMPORAL, se realizó un pequeño estudio mediante un cuestionario compuesto íntegramente por preguntas de MAPPING en las que se utilizaban *imágenes respuesta* que partían de la misma imagen original

La reutilización de imágenes fomenta el recuerdo de las mismas.

con diferentes métodos de procesamiento de imágenes. En todas las preguntas de MAPPING se solicitaba a los encuestados que marcaran dónde consideraban que se encontraba un cierto objeto. Se pudo observar que la precisión espacial de los entrevistados mejoraba conforme iban respondiendo a más preguntas que utilizaban la misma imagen original, casi independientemente del grado de calidad de la *imagen respuesta*. Es decir, al ir respondiendo preguntas, las respuestas iban aumentando la precisión espacial, siempre que el método de extracción de bordes obtuviese algunos contornos del objeto que se buscaba, sin necesidad de que el contorno del objeto se hubiese obtenido completamente. Se pudo observar un comportamiento interesante y es que en cuanto los entrevistados contestaban a alguna pregunta en la que el contorno del objeto estuviese completo, en todas las preguntas posteriores la precisión espacial de la respuesta era muy elevada. Se puede decir, que en cuanto conseguían encontrar el objeto por primera vez, ya lo encontraban siempre.

Más evidente fue el análisis del tiempo de respuesta, en el que se pudo observar como dicho tiempo de respuesta se reducía al responder más preguntas, llegando a ser casi constante en las últimas, prácticamente sin influencia de la calidad de la *imagen respuesta* que se mostrase. Se puede afirmar que los entrevistados recordaban la zona de la imagen dónde se encontraba el objeto y no tenían que volver a buscarlo por toda la imagen. También, se puede indicar que el proceso de búsqueda y detección se realizaba cada vez de manera más automática por parte de los encuestados tras responder a más preguntas que se basan la misma imagen original.

Por tanto, si se reutilizan imágenes en preguntas con respuestas de MAPEADO ESPACIAL Y TEMPORAL, ¿cuántas preguntas con otras imágenes diferentes hay introducir entre medio para que los entrevistados olviden la imagen? Para responder a esta pregunta, se realizó un experimento que determinó que, dos preguntas de MAPPING que se basen en la misma imagen original deben estar separadas, al menos, por otras 3 preguntas y, a ser posible, que dichas preguntas no sean del mismo tipo, sino que sean otros diferentes (SELECCIÓN, PUNTUACIÓN, etc.).

8.4.2 Uso de imágenes reales

El uso de imágenes reales permite a los entrevistados comparar cada *imagen respuesta* con la imagen real a partir de la cual se han generado. Utilizando las imágenes reales, los entrevistados obtienen mayor información visual, lo que permite en muchos casos eliminar algunas "dudas". Sin embargo, al presentar las imágenes reales, existe un elevado riesgo de introducir *sesgo por finalidad*. Esto es debido a que el entrevistado deja de comparar únicamente las *imágenes respuesta* entre sí, para primero comparar cada una de ellas con la imagen real y, después, comparar las *imágenes respuesta* entre sí. Mediante este proceso de comparación en dos pasos, el entrevistado, de manera consciente o no, tiende a RELLENAR zonas sensibles en las *imágenes respuesta* que, en caso de no tener la imagen original, descartaría (o aceptaría, dependiendo de cada entrevistado).

En pruebas preliminares realizadas, se ha podido determinar que exis-

El tiempo de búsqueda en preguntas de MAPPING con imágenes repetidas se va reduciendo conforme se contestan más preguntas con la misma imagen original.

te una cierta influencia en la respuesta de los entrevistados según la presencia o no de la imagen original, es decir, que los entrevistados respondían de forma ligeramente diferente ante las mismas *imágenes respuesta* dependiendo de si se incluía la imagen original en la pregunta. En particular, se realizó un experimento con dos cuestionarios sobre un grupo de entrevistados en los que en el primer cuestionario se les solicitaba que puntuasen la calidad de un grupo de *imágenes respuesta* presentando para cada pregunta en dicho cuestionario las imágenes originales. En el segundo cuestionario, se les solicitaba que puntuasen las mismas *imágenes respuesta* sin mostrarles las imágenes originales. Este segundo cuestionario, se les presentó varios días después para fomentar el *olvido* de las imágenes por parte de los entrevistados y alterando el orden de las preguntas. Los cambios en las respuestas de los encuestados entre los dos cuestionarios se produjeron generalmente en las *imágenes respuesta* de calidad intermedia, ya que las de alta calidad, siguieron obteniendo altas puntuaciones; al igual que las de baja calidad, que siguieron puntuándose de manera muy baja. Sin embargo, en las *imágenes respuesta* de calidad intermedia se observó un comportamiento no generalizable, había entrevistados que aumentaban la puntuación asignada anteriormente, mientras que otros entrevistados redujeron la puntuación de la misma imagen. Por tanto, se podría afirmar que el uso de imágenes reales ayuda a los entrevistados a decidir con mayor información visual, pero que dicha información visual no es tratada de igual manera por todos los usuarios y por lo tanto, el comportamiento no es reproducible en todos los casos.

8.4.3 Número de imágenes respuesta por pregunta

¿Cuántas *imágenes respuesta* como máximo se deberían utilizar en cada pregunta? Para contestar a esta pregunta se realizó una encuesta de pilotaje con un grupo cerrado de 6 encuestados con un grado de VISIÓN ESTÁNDAR y 1 con BAJA VISIÓN por cataratas. En dicha encuesta se pusieron diferentes preguntas con distinto número de *imágenes respuesta*. Se incluyeron preguntas de PUNTUACIÓN y SELECCIÓN, tanto SIN REPETICIÓN como CON REPETICIÓN. Tras contestar la encuesta, a los encuestados se les realizó una entrevista personal en la que se les preguntó por el número de *imágenes respuesta* e indicaron que dependía del tipo de pregunta. La apreciación de la mayoría (incluyendo la respuesta de la persona con BAJA VISIÓN) fue:

- Preguntas de PUNTUACIÓN (CON O SIN REPETICIÓN): Menos de 8 imágenes.
- Preguntas de SELECCIÓN SIN REPETICIÓN: Menos de 4 imágenes.
- Preguntas de SELECCIÓN CON REPETICIÓN: Entre 4 y 6 imágenes.
- Preguntas de MAPPING: No aplicable.

Los entrevistados indicaron, para cada caso, que más *imágenes respuesta* les exigía mucho esfuerzo para responder ya que tenían que comparar muchas imágenes entre sí. Por el contrario, con un menor número de ellas no había suficientes donde escoger y sentían un cierto grado de

frustración al tener que seleccionar imágenes con las que no estaban totalmente de acuerdo.

8.5 SISTEMA WEB PARA CUESTIONARIOS MULTIMEDIA

COMO ya se ha expuesto en el Capítulo 4, hay dos tipos de cuestionarios según el mecanismo de recolección de respuestas: *entrevistas estructuradas* y *cuestionarios auto-administrados*. Para la evaluación de la calidad de bordes, lo ideal sería realizar *entrevistas estructuradas*. Sin embargo, la gran dificultad que se presenta para la ubicación de pacientes con BAJA VISIÓN en un único sitio para poder realizar las entrevistas hace que este método no sea factible. Por el contrario, el mecanismo de *cuestionarios auto-administrados* presenta más virtudes para la evaluación de calidad de bordes, ya que permite que los usuarios puedan parar y continuar la evaluación según sus propios intereses.

Los pacientes con BAJA VISIÓN suelen presentar cansancio visual cuando realizan tareas visualmente exigentes.

De manera habitual, los *cuestionarios auto-administrados* se han realizado utilizando el soporte papel: se entrega un *cuadernillo* que contiene las preguntas y una o varias *hojas de respuesta*, en las que cada entrevistado se encarga de valorar las preguntas según unas determinadas escalas. En general, la cumplimentación de los cuestionarios es en muchos casos una tarea ardua para el entrevistado, ya que tiene que seleccionar unas respuestas a un determinado número de preguntas, pasar correctamente dichas respuestas a la hoja de respuestas para el procesamiento estadístico posterior por parte del encuestador, etc. Este proceso suele provocar el rechazo de los encuestados cuando son muchas preguntas o cuando el proceso de selección de respuestas es complejo, por lo tedioso de la tarea. Debido a todo esto, en muchos casos, los encuestados abandonan las encuestas sin contestarlas. Otro aspecto que involucra el uso del soporte papel y que hay que tener en cuenta es tanto la cantidad de papel como de tintas que se utiliza, en muchos casos para un único uso. Este uso desmesurado de papel y de tinta es muy poco ecológico, por su efectos contaminantes sobre el medio ambiente. Todos estos efectos se acentúan en cuestionarios con imágenes, ya que dichas imágenes han de imprimirse en alta calidad y en algún tipo de papel fotográfico o, en cualquier caso, de mayor grosor que el papel habitual, y con una gran cantidad de tinta de varios colores. Tampoco es posible el uso de papel reciclado, ya que los colores y la resolución se ven afectados en este tipo de papel. Todos estos aspectos elevan el coste final del cuestionario. Como ejemplo, en la Fig. 8.1 se muestra al autor de [47] fotografiado junto con una columna de papel impreso correspondiente con las 19600 páginas de todas las encuestas realizadas para dicho trabajo. Esta imagen muestra gráficamente de manera muy evidente el enorme gasto tanto en papel como en costes de impresión que conlleva una evaluación mediante cuestionarios impresos. Adicionalmente, a este tipo de evaluación hay que incorporar otros costes, como es el tiempo empleado en recopilar las respuestas y transcribirlas a una base de datos para su posterior procesamiento.

Frente a la aproximación basada en la impresión en papel de los cuestionarios, se encuentra la cumplimentación de los cuestionarios utilizando algún sistema Web. La alta calidad de los monitores, con elevadas resoluciones gráficas, y la disponibilidad de los cuestionarios en



Figura 8.1
Imagen extraída de
HEATH *et al.* [47],
muestra las 19600
hojas de evaluación
impresas de todas las
encuestas realizadas.

todo momento y en todo sitio con conexión a Internet, hacen que los sistemas Web se postulen como una buena solución para enfrentarse a la evaluación de encuestas de tipo multimedia. Por todo lo anterior, para proporcionar soporte a los cuestionarios para la evaluación de la calidad de distintas técnicas de procesamiento de imágenes se ha planificado un sistema Web.

Siguiendo las recomendaciones expuestas en este capítulo se ha desarrollado una herramienta que permite el diseño de cuestionarios para la evaluación de contenidos multimedia. A su vez, este sistema facilita la gestión de dichos cuestionarios tanto por parte del encuestador para su tratamiento, como de los encuestados para contestarlos. Finalmente, proporciona al encuestador la evaluación de los resultados de los cuestionarios tras aplicarles diversos factores de ponderación.

Esta herramienta ha sido implementada como un sistema Web, para permitir que los encuestados puedan acceder de manera remota a los cuestionarios. De manera adicional, los usuarios pueden interrumpir y retomar los cuestionarios en cualquier momento. Esto es una práctica bastante común en pacientes con problemas de visión debido al cansancio visual que les acarrea el esfuerzo de ver las imágenes.

8.5.1 Preguntas de PUNTUACIÓN

En la Fig. 8.2 se muestra un ejemplo de pregunta de PUNTUACIÓN de una encuesta de pilotaje para la evaluación de bordes.

Figura 8.2
Pregunta de PUNTUACIÓN CON REPETICIÓN.

Pregunta de Puntuación


Respondida:

[<<<] [<] 1 [2] 3 4 5 6 7 [>] [>>>]

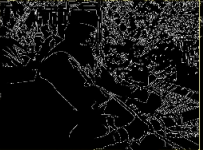
A continuación Vd. va a poder ver 4 imágenes de bordes, todas obtenidas utilizando el mismo mecanismo, pero cada una de ellas con un parámetro de selección de bordes diferente. Puntúe de 0 a 10 cada imagen según la calidad que Vd. considere que tiene dicha imagen según los bordes que en ellas se vean y lo que a Vd. estos bordes le ayuden para ver los distintos objetos de la imagen. **RECUERDE PINCHAR EN CADA IMAGEN PARA VERLAS EN SU TAMAÑO ORIGINAL.**

Escala de Calidad:

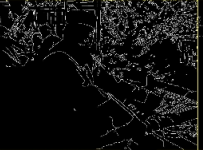
- 0 calidad pésima, no se sabe que representa la imagen
- 10 calidad perfecta, se han detectado los bordes de manera perfecta y gracias a ellos puedo ver los distintos objetos que componen la imagen




Puntuación:



Puntuación:



Puntuación:



Puntuación:

Nótese que debajo de cada *imagen respuesta* existe una caja de texto que admite un valor numérico. Los límites inferior y superior de dicho valor se puede parametrizar por parte del encuestador cuando construye el cuestionario, de manera que impida la introducción de valores fuera de rango.

8.5.2 Preguntas de SELECCIÓN SIN REPETICIÓN

En la Fig. 8.3 se muestra un ejemplo de pregunta de SELECCIÓN SIN REPETICIÓN de una encuesta de pilotaje para la evaluación de bordes.

Figura 8.3
Pregunta de SELECCIÓN SIN REPETICIÓN.

Pregunta de Selección sin repetición

Respondida:

[<<<] [<] 1 [2] 3 4 5 6 [>] [>>>]

Las siguientes 3 imágenes muestran los mapas de bordes generados utilizando en cada uno un método diferente. En la imagen original, que no se muestra, existe una motocicleta en sombras. Intente encontrar la motocicleta en cada una de las imágenes de bordes. De entre todas las imágenes que se presentan, seleccione una única imagen de bordes, en la que Vd. considere que es la de mejor calidad de entre todas las que sea capaz de detectar dicha motocicleta. **RECUERDE PINCHAR EN CADA IMAGEN PARA VERLAS EN SU TAMAÑO ORIGINAL.**







Debajo de cada *imagen respuesta* hay un botón circular que muestra si se ha seleccionado la imagen superior asociada y que, a su vez, sólo permite que se encuentre activo uno de dichos botones. De tal forma, que si ya hubiese un botón seleccionado y se escoge otro botón, el primero se desactivaría y pasaría a activarse el segundo.

8.5.3 Preguntas de SELECCIÓN CON REPETICIÓN

En la Fig. 8.4 se muestra un ejemplo de pregunta de SELECCIÓN CON REPETICIÓN de una encuesta de pilotaje para la evaluación de bordes.

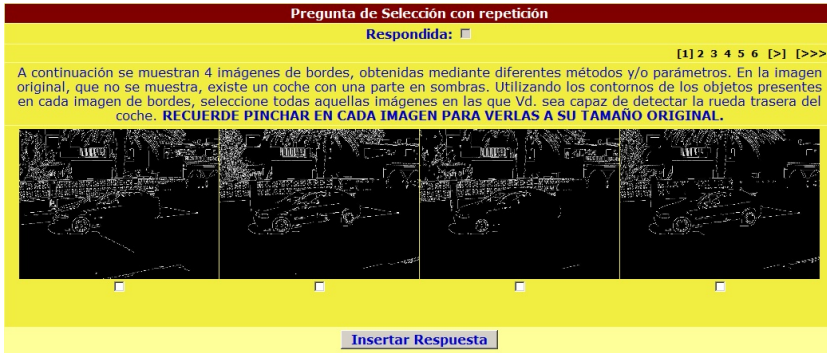


Figura 8.4
Pregunta de SELECCIÓN CON REPETICIÓN.

Este tipo de preguntas tienen asociadas a cada *imagen respuesta* un cuadro de verificación, que admite que varios de ellos se encuentren activados simultáneamente.

8.5.4 Preguntas de MAPEADO ESPACIAL Y TEMPORAL

Finalmente, en la Fig. 8.5 se muestra un ejemplo de pregunta con MAPEO ESPACIAL Y TEMPORAL de la respuesta para una encuesta de pilotaje para la evaluación de bordes.

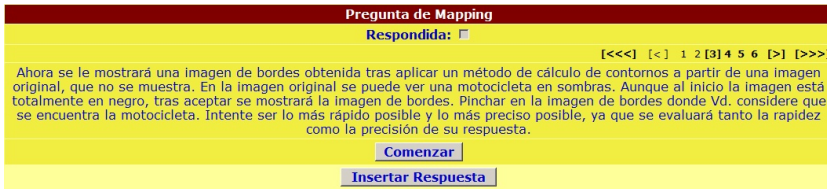


Figura 8.5
Pregunta de MAPPING.

En este tipo de pregunta se presentan 2 ventanas, la primera, mostrada en la Fig. 8.5, con el texto de la pregunta y un botón *Comenzar*, que al pulsarlo muestra la imagen de bordes que se pretende evaluar. Al contemplarse en este tipo de preguntas el tiempo de respuesta, la carga de la imagen se realiza en segundo plano mientras que se muestra el texto de la pregunta. Al pulsar el botón de comienzo, se inicia el contador de tiempo y se muestra la imagen precargada en una ventana externa (por lo que para un correcto funcionamiento del sistema hay que desactivar cualquier tipo de software bloqueador de *pop-up*). En dicha ventana, se muestra la imagen y se espera a que el usuario clique con el ratón en algún sitio de la ventana, momento en el que el contador de tiempos se para y se pregunta si se desea almacenar la posición marcada como respuesta. En caso afirmativo, se almacena en el sistema tanto el tiempo empleado en clicar como la posición (x, y) de dicha respuesta.

8.6 CONCLUSIONES

En este capítulo se ha analizado el llamado *sesgo por finalidad*, que ha sido identificado como una fuente de distorsión en las respuestas de los encuestados y que no se tiene conocimiento de que hubiera sido descrito previamente. También, se han propuesto algunas recomendaciones de diseño de cuestionarios que incorporen elementos visuales dentro de las opciones de respuesta. Estas recomendaciones se han obtenido mediante experimentación y metaencuestas sobre entrevistados que cumplieron encuestas de pilotaje. Se ha propuesto un conjunto de tipos de preguntas que facilitan el diseño correcto de este tipo de cuestionarios. Hay que hacer especial hincapié en las preguntas con respuestas de MAPEADO ESPACIAL Y TEMPORAL, ya que son una aportación muy importante que hace esta Tesis para el desarrollo de cuestionarios para evaluación de imágenes.

A pesar de que estas recomendaciones son, en muchos casos de sentido común, sería conveniente realizar un estudio estadístico más pormenorizado que corroborase científicamente las afirmaciones que se han realizado en este capítulo. Este estudio no se ha realizado porque excedía los límites de la presente Tesis, pero desde luego abre un interesante campo de trabajo para futuras investigaciones y tesis doctorales.

También es de sumo interés, y es otra de las aportaciones importantes de esta Tesis Doctoral, la construcción de un sistema Web para la construcción de cuestionarios con contenidos multimedia por parte de los encuestadores, y para el manejo y relleno de dichas encuestas por parte de los encuestados.

8.6.1 Conclusiones de la Parte II

Con este capítulo se finaliza la Parte II, *Método y Materiales*, en la que se han descrito las aportaciones científicas más relevantes de esta Tesis Doctoral. A lo largo del Capítulo 6 se describió matemáticamente el nuevo operador *Convolución-LIP*, exponiéndose además dos implementaciones diferentes del mismo, *DGConv-LIP* y *FGConv-LIP*. En el Capítulo 7 se especificaron los criterios generales y específicos que se han seleccionado para este trabajo y que han condicionado en gran medida las decisiones de investigación tomadas. Para evaluar los resultados de los algoritmos de extracción de bordes diseñados bajo el modelo LIP utilizando el operador *Convolución-LIP*, en el Capítulo 8 se han presentado un conjunto de recomendaciones para el diseño de cuestionarios de evaluación de calidad de bordes. También se ha presentado una herramienta Web que ha sido diseñada expresamente para el diseño y evaluación de los cuestionarios de contenidos multimedia. El compendio de capítulos que forman la Parte II es el núcleo central de esta Tesis Doctoral, ya que en estos capítulos se han expuesto las innovaciones científicas. La siguiente Parte III, *Resultados Experimentales*, comenzará con un capítulo en el que se presentarán experimentos y aplicaciones del nuevo operador propuesto *Convolución-LIP*, comparando los tiempos de computación entre diferentes implementaciones. Tras ese capítulo, se mostrarán los resultados obtenidos a partir del

cuestionario de evaluación de imágenes de bordes, tanto para pacientes con BAJA VISIÓN como para pacientes con un grado de VISIÓN ESTÁNDAR.

Parte III

RESULTADOS EXPERIMENTALES

ÍNDICE

9.1	LIP-Sobel	158
9.1.1	Formulación	158
9.1.2	Resultados experimentales	159
9.2	Filtros de emborronamiento	161
9.2.1	Filtro Media: LIP-Avg	162
9.2.2	Filtro Gaussiano: LIP-GBLur	163
9.3	Filtros para la extracción de bordes	165
9.3.1	Laplaciano de Gaussianas: LIP-LoG	165
9.3.2	LIP-Canny	169
9.4	Tiempos de cómputo	171
9.4.1	Aumento del rendimiento (<i>Speedup</i>)	172
9.4.2	LIP-Sobel	172
9.4.3	Algoritmos de emborronamiento: LIP-Avg, LIP-GBLur	173
9.4.4	Extracción de bordes: LIP-LoG	174
9.4.5	Extracción de bordes: LIP-Canny	175
9.5	Resumen de resultados	176
9.5.1	LIP-Sobel	177
9.5.2	LIP-Avg, LIP-GBLur	178
9.5.3	LIP-LoG	179
9.5.4	LIP-Canny	180
9.6	Conclusiones	180

Para mostrar la aplicabilidad de la *Convolución-LIP* se mostrarán una serie de experimentos con diversas técnicas habituales y bien conocidas en la literatura científica tradicional, adaptadas al paradigma LIP. Se han realizado experimentos con algoritmos de extracción de bordes y otros algoritmos de filtrado paso baja. Se han realizado experimentos con todos los algoritmos seleccionados en el Cap. 7. Se comienza con un algoritmo ya implementado en LIP, el método *LIP-Sobel*. Éste muestra que se obtienen los mismos resultados implementando el algoritmo con el método original y utilizando el mecanismo de *Convolución-LIP*, implementada ésta mediante cualquiera de las dos versiones (DGLIP-CONV o FGLIP-CONV). Tras este primer experimento, se realizan experimentos con filtros de emborronamiento para la reducción del ruido, basados en un filtrado *media* y en un filtrado gaussiano. Se continuará con experimentos con filtros de extracción de bordes, el *Laplaciano de Gaussianas* y el método de *Canny*. Para finalizar el capítulo se muestran las comparativas de tiempos de cómputo y las ganancias de velocidad, en las que se puede ver que FGLIP-CONV es la implementación más rápida para todos los algoritmos que se diseñen bajo el modelo LIP.

9.1 LIP-SOBEL

Inicialmente, se va a mostrar que los algoritmos diseñados utilizando cualquiera de las dos implementaciones del operador *Convolución*-LIP (DGLIP-CONV o FGLIP-CONV), definidas en el Cap. 6, producen el mismo resultado que los algoritmos previamente diseñados por otros autores siguiendo la metodología LIP. Para ello, la comparación se realizará con un método llamado *LIP-Sobel*, propuesto inicialmente por Deng y Pinoli [29]. Este método es la reformulación del método de extracción de bordes de *Sobel* siguiendo el modelo LIP. Los autores diseñaron unas fórmulas adaptadas a LIP, mediante un proceso totalmente *off-line*, a partir de los filtros habituales de *Sobel*, ya que los valores de los filtros que iban a aplicar eran constantes conocidas y fijadas previamente. Estas fórmulas diseñadas son específicas para este proceso en particular y no puede exportarse a otras aplicaciones o algoritmos, por lo tanto se puede afirmar que no es un proceso general, sino específico para ese método. Sin embargo, es fácilmente deducible que la metodología sí es generalizable a otros filtros, siempre y cuando los valores de dichos filtros sean constantes y conocidos *a priori*.

9.1.1 Formulación

Ya que el filtro de *convolución* es suficientemente pequeño se puede asumir que la *iluminación* dentro de un área tan reducida es constante. Esto ha producido un método que se comporta de manera casi invariante ante los cambios de iluminación y proporciona una respuesta mucho más satisfactoria para la detección de los contornos de objetos en sombras que el método *Sobel* original (es decir, sin LIP). Este nuevo método detecta bordes tanto en zonas bien iluminadas como en las pobremente iluminadas o en las sobreiluminadas.

Sea un área o un entorno de *tonos de gris*, definida por:

$$\begin{array}{ccc} \hat{f}_1 & \hat{f}_2 & \hat{f}_3 \\ \hat{f}_4 & \hat{f}_5 & \hat{f}_6 \\ \hat{f}_7 & \hat{f}_8 & \hat{f}_9 \end{array} \quad (9.1)$$

donde \hat{f}_5 es el píxel central del entorno que se está considerando.

Utilizando (2.34), en la pág. 38, que muestra las máscaras originales del método de *Sobel*, y siguiendo el método directo de LIP, se ha definido el VECTOR de *tonos de gris* del método de *LIP-Sobel*, $\mathbf{g} = (g_x, g_y)$, como:

$$g_x = \left(\hat{f}_1 \Delta \left(2\Delta \hat{f}_4 \right) \Delta \hat{f}_7 \right) \Delta \left(\hat{f}_3 \Delta \left(2\Delta \hat{f}_6 \right) \Delta \hat{f}_9 \right) \quad (9.2)$$

$$g_y = \left(\hat{f}_1 \Delta \left(2\Delta \hat{f}_2 \right) \Delta \hat{f}_3 \right) \Delta \left(\hat{f}_7 \Delta \left(2\Delta \hat{f}_8 \right) \Delta \hat{f}_9 \right) \quad (9.3)$$

Los autores pudieron definir las fórmulas que generan el vector del método *LIP-Sobel* porque se conocen los valores del filtro previamente y además, son constantes (valores positivos y negativos, que se tradujeron a la formulación directa de LIP mediante Δ ó $\bar{\Delta}$, respectivamente).



Figura 9.1
Imagen peppers original.

Utilizando la función isomórfica de transformación, $\varphi(\cdot)$, descrita en (2.16), se obtiene (ver demostración en [29, Apéndice 3]):

$$g_x = M - M \left(\frac{f_3 f_6^2 f_9}{f_1 f_4^2 f_7} \right) \quad (9.4)$$

$$g_y = M - M \left(\frac{f_1 f_2^2 f_3}{f_7 f_8^2 f_9} \right) \quad (9.5)$$

Se deduce fácilmente que (9.4) y (9.5) se pueden obtener directamente de la formulación de DGLIP-CONV, mostrada en (6.11), pág. 108, utilizando las máscaras definidas en (2.34)). Por tanto, DGLIP-CONV es una formulación más general que el método planteado por Deng y Pinoli, y consecuentemente, al ser equivalentes las dos formulaciones, FGLIP-CONV, también es una formulación más general que la propuesta por Deng y Pinoli.

9.1.2 Resultados experimentales

Aunque en el Cap. 7 se han seleccionado una serie de imágenes, para comprobar el correcto funcionamiento del método LIP-Sobel en sus diferentes implementaciones y facilitar la comparación visual, se ha decidido utilizar alguna imagen conocida y habitual en las comparaciones del ámbito del PROCESAMIENTO DE IMÁGENES. Para ello, se ha seleccionado la conocida imagen peppers (ver Fig. 9.1).

La Fig. 9.2 muestra la imagen resultante tras aplicar un proceso de oscurecimiento no lineal sobre la imagen original (Fig. 9.1). El proceso de oscurecimiento se ha realizado con la fórmula:

$$\text{dark}(x, y) = \text{img}(x, y) \cdot \left(0.1 + \frac{5 \cdot \sin(2\pi x)}{6 \cdot \text{width}} \right). \quad (9.6)$$

Donde “width” es la anchura de la imagen que se desea oscurecer e $\text{img}(x, y)$ es el valor del píxel (x, y) de dicha imagen original.

A continuación, se muestran los resultados de LIP-Sobel construidos con tres metodologías diferentes (siguiendo el método propuesto por

Figura 9.2
Imagen peppers oscu-
recida.



Figura 9.3
Sobel estándar apli-
cado sobre la imagen
peppers oscurecida.

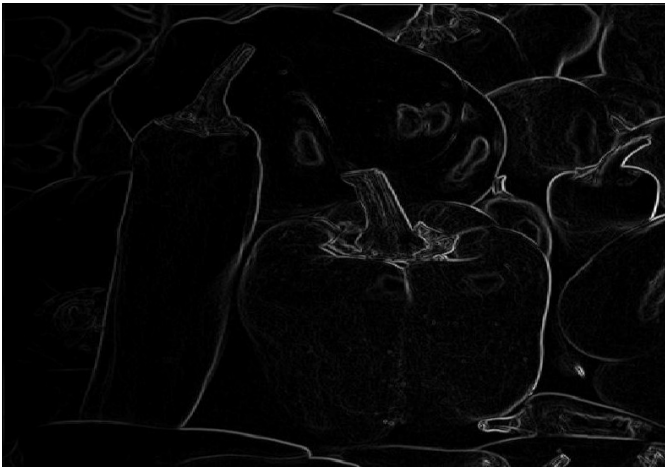
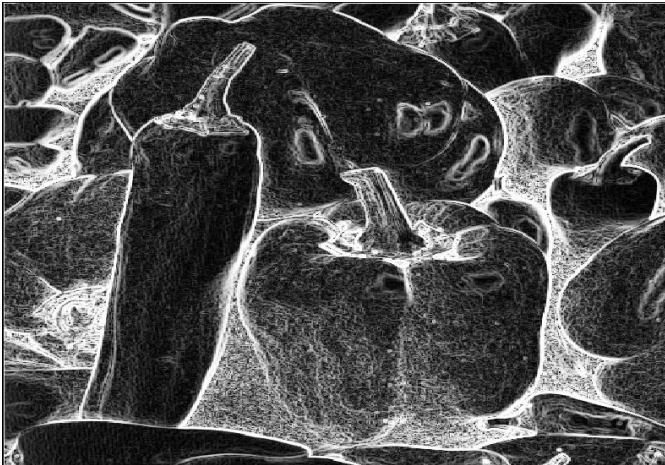


Figura 9.4
LIP-Sobel aplica-
do sobre la imagen
peppers oscurecida.



MSE: Del inglés,
MEAN SQUARED
ERROR, Error
cuadrático medio.

Deng y Pinoli [29], mediante DGLIP-Conv y mediante FGLIP-Conv). Como era de esperar, los resultados obtenidos por los tres métodos son exactamente iguales, como demuestra que el MSE sea cero en todos los casos.

LIP-Sobel implementado mediante el procedimiento de Deng y Pinoli

En la Fig. 9.3 se muestra la imagen resultante de aplicar el método *Sobel* estándar sobre la imagen oscurecida (Fig. 9.2). Se ha aplicado sobre la Fig. 9.2 el método *LIP-Sobel* tal y como fue propuesto por Deng y Pinoli, descrito en (9.4) y (9.5). En la Fig. 9.4 se muestra la magnitud del gradiente de la imagen resultante. Como se puede observar en la imagen de la Fig. 9.3, que muestra la aplicación del método *Sobel* original, los bordes presentes en las regiones oscuras de la imagen original no se han detectado con claridad. Frente a esta imagen, se puede observar la Fig. 9.4, donde se muestra el resultado del método *LIP-Sobel*, en la que los bordes en las regiones oscuras se detectan de manera más homogénea. Este comportamiento es un beneficio colateral, debido a que hace que la componente de *iluminación* se pueda considerar constante en un entorno pequeño y gracias al comportamiento logarítmico de LIP, esta componente se elimina automáticamente al realizar la operación *Convolución-LIP* con un filtro cuya suma de componentes sea nula, o bien, al realizar operaciones Δ y Δ que se compensen entre sí.

LIP-Sobel implementado mediante FGLIP-CONV y DGLIP-CONV

El filtro *Sobel* estándar, mostrado en (2.34), es un filtro separable, compuesto por dos vectores:

$$\begin{aligned} \mathbf{a} &= [-1, 0, 1] \\ \mathbf{b} &= [1, 2, 1] \end{aligned} \quad (9.7)$$

Usando la propiedad de separabilidad de los filtros de *Sobel*, se puede usar el operador *Convolución-LIP*, implementado tanto por el método DGLIP-CONV como por el método FGLIP-CONV. De modo que para realizar el algoritmo *LIP-Sobel* mediante el método DGLIP-CONV, se aplica (6.11) utilizando \mathbf{a} y \mathbf{b} tal y como se ha propuesto en (9.7) a la imagen oscurecida (Fig. 9.2). De manera análoga, se obtiene el mismo resultado utilizando el algoritmo *LIP-Sobel* mediante el mecanismo FGLIP-CONV, para lo que se aplicará (6.13) con \mathbf{a} y \mathbf{b} , descritos en (9.7). Los bordes detectados por *LIP-Sobel* mediante cualquiera de los tres procedimientos (Deng y Pinoli, DGLIP-CONV y FGLIP-CONV) son los mismos y, como se ha comentado con anterioridad, el MSE es cero.

9.2 FILTROS DE EMBORRONAMIENTO

En la aplicación previa se ha mostrado que los dos mecanismos propuestos para el desarrollo del operador *Convolución-LIP* proporcionan los mismos resultados para un algoritmo en particular mediante el uso de un algoritmo diseñado específicamente. Sin embargo, el nuevo operador *Convolución-LIP* es más interesante y versátil, y a continuación, se mostrarán experimentos que van a permitir refrendar esta afirmación.

Siguiendo el mismo procedimiento que utilizaron Deng y Pinoli para obtener la formulación con la que se implementó el método *LIP-Sobel*

original, se podría diseñar cualquier otro filtro. Sin embargo, cada filtro debería calcularse de manera explícita previamente; por ejemplo, aunque simplemente se deseara modificar el tamaño del filtro para que tenga un entorno diferente, se debería calcular un filtro completamente nuevo, sin poder adaptarse el calculado previamente. Justo al contrario, el operador *Convolución-LIP*, ya sea implementado mediante DGLIP-CONV o mediante FGLIP-CONV, permiten utilizar el mismo código, simplemente cambiando los parámetros de entrada (los filtros 1D utilizados), lo cual se traduce en una ganancia en flexibilidad y, a su vez, en robustez de las aplicaciones. Además de estas ventajas, se obtiene una mejora en el tiempo de desarrollo y sobre todo en el tiempo de cómputo, como se mostrará en la Sec. 9.4.

9.2.1 Filtro Media: LIP-Avg

En este apartado, se va a mostrar el que se puede considerar como “mejor caso” en términos de velocidad para el mecanismo de construcción de filtros siguiendo la filosofía LIP desarrollada por Deng y Pinoli. Para ello, sea un filtro en el que todos los elementos son constantes y con valores idénticos, por ejemplo, un filtro *media*. Se afirma que éste es el “mejor caso” para el mecanismo de Deng y Pinoli, ya que la potenciación, que es una operación computacionalmente muy costosa, se puede realizar solo una única vez por filtro, tal y como se demuestra a continuación. Se compara este resultado con las imágenes resultantes obtenidas mediante DGLIP-CONV y mediante FGLIP-CONV, comprobándose que el resultado es el mismo, pero el tiempo de ejecución de estos últimos es menor que el “mejor caso” del mecanismo propuesto por Deng y Pinoli.

LIP-Avg implementado mediante el procedimiento de Deng y Pinoli

Se define un filtro de tamaño $n \times n$ píxeles para que calcule la media de dicho entorno siguiendo el paradigma LIP. Como primer paso, se fija el valor de n , para poder construir la fórmula específica asociada al filtro *media*. Por ejemplo, si se toma $n = 3$, se puede utilizar la misma definición de entorno de un píxel que en (9.1), con lo que se obtiene:

$$\text{avg}_{\text{LIP}}^{3 \times 3}(f) = \left(\begin{array}{ccc} \left(\frac{1}{9} \Delta \hat{f}_1\right) \Delta \left(\frac{1}{9} \Delta \hat{f}_2\right) \Delta \left(\frac{1}{9} \Delta \hat{f}_3\right) & \Delta & \\ \left(\frac{1}{9} \Delta \hat{f}_4\right) \Delta \left(\frac{1}{9} \Delta \hat{f}_5\right) \Delta \left(\frac{1}{9} \Delta \hat{f}_6\right) & \Delta & \\ \left(\frac{1}{9} \Delta \hat{f}_7\right) \Delta \left(\frac{1}{9} \Delta \hat{f}_8\right) \Delta \left(\frac{1}{9} \Delta \hat{f}_9\right) & & \end{array} \right) \quad (9.8)$$

Operando sobre esta fórmula, se obtiene:

$$\text{avg}_{\text{LIP}}^{3 \times 3}(f) = M - (f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \cdot f_6 \cdot f_7 \cdot f_8 \cdot f_9)^{\frac{1}{9}} \quad (9.9)$$

Que se podría abreviar como,

$$\text{avg}_{\text{LIP}}^{3 \times 3}(f) = M - \left(\prod_{i=1}^9 f_i \right)^{\frac{1}{9}} \quad (9.10)$$

Hay que notar que se puede tomar un exponente común ($\frac{1}{9}$). Al utilizarlo, el tiempo de ejecución puede disminuir, ya que se reduce sensiblemente la complejidad de la fórmula que se ha de codificar.

También se ha probado con un entorno de mayor tamaño: 5×5 . Para este caso, la nueva fórmula obtenida es:

$$\text{avg}_{\text{LIP}}^{5 \times 5}(f) = M - \left(\prod_{i=1}^{25} f_i \right)^{\frac{1}{25}} \quad (9.11)$$

Está claro que para cada valor diferente de n , es decir, para cada tamaño del entorno que se considere, se tiene que calcular una nueva fórmula y, por lo tanto, es necesario reescribir nuevo código para cada nuevo tamaño del filtro.

LIP-Avg implementado mediante DGLIP-CONV y FGLIP-CONV

Mediante el operador *Convolución-LIP* se puede construir fácilmente el filtro media siguiendo la metodología LIP para entornos de cualquier tamaño. Para implementar *LIP-Avg* de 3×3 píxeles, sólo se tiene que aplicar (6.11) ó (6.13), dependiendo de si la implementación es siguiendo el método DGLIP-CONV o FGLIP-CONV respectivamente, y utilizando como componentes del filtro los vectores $\mathbf{a} = \mathbf{b} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. Para calcular *LIP-Avg* de 5×5 , se utiliza (6.11) ó (6.13) con los vectores $\mathbf{a} = \mathbf{b} = [\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}]$. Se puede deducir fácilmente que, simplemente, cambiando los filtros que se utilizan como parámetros de entrada, se puede obtener el cálculo de cualquier filtro *media* para diferentes tamaños del entorno, sin necesidad de escribir una sola línea de código adicional.

9.2.2 Filtro Gaussiano: LIP-GBlur

En el apartado anterior, se han mostrado un par de ejemplos de lo que se puede considerar el caso más favorable para la metodología propuesta por Deng y Pinoli. Este “mejor caso” se debe a la reducción del número de exponenciaciones utilizadas, ya que cada píxel del entorno se eleva al mismo valor, y por tanto, se puede tomar un exponente común. Por tanto, el “peor caso” para la metodología de Deng y Pinoli es un filtro con valores diferentes, en el que no se pudiera optimizar sacando un exponente común. Un ejemplo de este tipo es el de emborronamiento gaussiano, donde el filtro está generado por una gaussiana. Se ha escogido un filtro gaussiano bidimensional puesto que es separable: la manera habitual de calcular un filtro gaussiano bidimensional es mediante la multiplicación matricial de un filtro gaussiano unidimensional (vector columna) por sí mismo traspuesto (vector fila). Para mostrar un ejemplo específico, se ha generado un filtro gaussiano de tamaño 7×7 (ver Tab. 9.1), realizado^o mediante la aproximación de una gaussiana (con $\sigma = 1.0$) para una malla de 7×7 píxeles.

A continuación, se muestran los resultados obtenidos por el filtro *LIP-GBlur*, que muestra el emborronamiento gaussiano siguiendo la filosofía LIP, mediante los tres métodos indicados anteriormente. Como era de esperar, no se han obtenido diferencias entre los resultados obte-

Tabla 9.1
Malla de 7×7 píxeles
de un filtro gaus-
siano bidimensional
(con $\sigma = 1.0$).

0.0001	0.0015	0.0067	0.0111	0.0067	0.0015	0.0001
0.0015	0.0183	0.0821	0.1353	0.0821	0.0183	0.0015
0.0067	0.0821	0.3679	0.6065	0.3679	0.0821	0.0067
0.0111	0.1353	0.6065	1.0000	0.6065	0.1353	0.0111
0.0067	0.0821	0.3679	0.6065	0.3679	0.0821	0.0067
0.0015	0.0183	0.0821	0.1353	0.0821	0.0183	0.0015
0.0001	0.0015	0.0067	0.0111	0.0067	0.0015	0.0001

nidos por los diferentes métodos para la implementación del mismo filtro.

LIP-GBLur implementado mediante el procedimiento de Deng y Pí-noli

Se ha extendido la definición del entorno de píxeles en el mismo sentido de (9.1) a un entorno de 7×7 píxeles. El filtro gaussiano utilizado se detalla en Tab. 9.1.

La fórmula que se aplica para obtener el emborronamiento gaussiano bajo el paradigma LIP, *LIP-GBLur*, se muestra en (9.12), aunque, por simplicidad, no se muestran todos los valores de la fórmula.

$$\begin{aligned}
 \text{GBLur}_{\text{LIP}}(f) = & \\
 & \left(\begin{array}{ccccccc}
 (0.0001 \triangle \widehat{f}_1) \triangle & (0.0015 \triangle \widehat{f}_2) \triangle & (0.0067 \triangle \widehat{f}_3) \triangle & & \dots & & \\
 \dots & (0.1353 \triangle \widehat{f}_{23}) \triangle & (0.6065 \triangle \widehat{f}_{24}) \triangle & (1.0000 \triangle \widehat{f}_{25}) \triangle & & \dots & \\
 \dots & & (0.0067 \triangle \widehat{f}_{47}) \triangle & (0.0015 \triangle \widehat{f}_{48}) \triangle & (0.0001 \triangle \widehat{f}_{49}) & &
 \end{array} \right) \quad (9.12)
 \end{aligned}$$

Análogamente a lo realizado para obtener (9.4) y (9.5), se obtiene la fórmula (9.13).

$$\text{GBLur}_{\text{LIP}}(f) = M - \left(\frac{f_1^{0.0001} \cdot \dots \cdot f_{24}^{0.6065} \cdot f_{25}^{1.0000} \cdot \dots \cdot f_{49}^{0.0001}}{M^{0.0001 \cdot \dots \cdot 0.6065 \cdot 1.0000 \cdot \dots \cdot 0.0001}} \right) \quad (9.13)$$

LIP-GBLur implementado mediante DGLIP-CONV y FGLIP-CONV

Como se ha indicado anteriormente, la matriz bidimensional de una gaussiana es un filtro separable, construido a partir de la multiplicación de un vector de una gaussiana unidimensional consigo mismo. El tamaño del filtro depende del valor del parámetro σ : cuanto mayor es σ , más grande debe ser el tamaño del filtro para que pueda tener suficiente precisión. Como regla general, la longitud del filtro gaussiano, centrado sobre una posición específica, debe ser de $3.5 \cdot \sigma^2$ valores a izquierda y derecha de dicha posición (está ampliamente aceptado el redondeo hacia el valor inferior). Por ejemplo, para $\sigma = 1$, se debe tomar un filtro gaussiano de tamaño 7. Como ejemplo, para dicho valor de σ , en los experimentos se han utilizado los siguientes vectores:

$$\mathbf{a} = \mathbf{b} = [0.011, 0.135, 0.606, 1, 0.606, 0.135, 0.011] \quad (9.14)$$

Es evidente, que para un valor más grande de σ , se deberá utilizar un filtro de mayor tamaño. Sin embargo, no hay necesidad alguna de reescribir ninguna línea de código, simplemente basta con cambiar los parámetros de entrada para obtener la salida deseada. Esto es una de las mayores contribuciones de esta nueva reformulación del operador *convolución* utilizando el paradigma LIP, propuesto en este trabajo.

9.3 FILTROS PARA LA EXTRACCIÓN DE BORDES

Como se ha expuesto en capítulos anteriores, los algoritmos de extracción de bordes serán las aplicaciones principales que se van a desarrollar en esta Tesis Doctoral. Por ese motivo, en este capítulo se mostrará el funcionamiento de los algoritmos de extracción de bordes seleccionados en el Cap. 7. El primero de los algoritmos de extracción de bordes seleccionados, *LIP-Sobel* ya se ha expuesto en la Sec. 9.1. A continuación, se describirán y mostrarán los dos restantes, *LIP-LoG* y *LIP-Canny*, que son la implementación bajo el modelo LIP de los algoritmos *LoG* y *Canny*, descritos en el Cap. 2.

9.3.1 Laplaciano de Gaussianas: LIP-LoG

Existen muchos métodos para la extracción de bordes de una imagen digital, como se observó anteriormente, *Sobel* es un algoritmo para la obtención de los contornos. Este método se basa en filtros fijos de un tamaño pequeño. Sin embargo, existen otros métodos que tienen filtros de tamaño variable y con valores en función del tamaño. Uno de estos métodos es el llamado *Laplaciano de Gaussianas*, también conocido como *LoG*. El operador *LoG* es la suma de la segunda derivada parcial en cada dirección sobre la imagen que previamente ha sufrido un emborronamiento gaussiano, aunque estas dos operaciones se hacen simultáneamente, resultando en una mayor efectividad. El operador bidimensional *LoG* se realiza mediante la *convolución* de una imagen con un filtro *LoG* bidimensional, tal y como se muestra en (9.15).

$$\text{LoG}(I, \sigma) = \nabla^2 \left\{ \text{conv2D}(I, G_\sigma(x, y)) \right\} = \text{conv2D} \left(I, \left(\nabla^2 G_\sigma(x, y) \right) \right) \quad (9.15)$$

donde $G_\sigma(x, y)$ simboliza la función 2D gaussiana descrita en (9.19) y $\nabla^2 G(x, y)$ es la suma de las segundas derivadas parciales direccionales de $G(x, y)$ mostradas en (9.16).

$$H(x, y) = \nabla^2 G(x, y) = \frac{\partial^2 G(x, y)}{\partial x^2} + \frac{\partial^2 G(x, y)}{\partial y^2} = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (9.16)$$

El detector de bordes *Laplaciano de Gaussianas* habitualmente se implementa utilizando un filtro bidimensional, como se ha indicado anteriormente; sin embargo, se ha demostrado que el filtro *LoG* es separable (ver [114] para la demostración de separabilidad del filtro *LoG*) y

dicha aproximación es la utilizada en este trabajo, tanto para implementar el filtro original *LoG* como para la versión *LIP-LoG* mediante *FGLIP-Conv*. Sin embargo, para implementar *LIP-LoG* utilizando la metodología propuesta por **Deng y Pinoli** no se pueden utilizar filtros separables ya que no ha sido desarrollado así originalmente; al contrario se ha utilizado un filtro bidimensional *LoG* descrito en (9.16). La formulación separable del filtro *LoG* se muestra en (9.17):

$$\begin{aligned} \text{LoG}(I, \sigma) = \nabla^2 \left\{ \text{conv2D}(I, G_\sigma) \right\} = & \text{conv1D} \left(\text{conv1D}(I, \mathbf{a}_\sigma(x)), \mathbf{g}_\sigma(y) \right) + \\ & \text{conv1D} \left(\text{conv1D}(I, \mathbf{a}_\sigma(y)), \mathbf{g}_\sigma(x) \right) \end{aligned} \quad (9.17)$$

donde $G_\sigma(x, y)$ es una función gaussiana bidimensional, $\mathbf{g}_\sigma(t)$ representa una función gaussiana unidimensional y $\mathbf{a}_\sigma(t)$ muestra la segunda derivada de la función gaussiana unidimensional, definidas respectivamente como:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (9.18)$$

$$\mathbf{g}_\sigma(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \quad (9.19)$$

$$\mathbf{a}_\sigma(t) = \frac{\partial^2 G_\sigma(t)}{\partial t^2} \quad (9.20)$$

El tamaño del filtro de *Sobel* es de 3×3 , sin embargo, para poder utilizar un filtro *LoG* se necesita un filtro con un tamaño mayor, para este caso en particular se utiliza un filtro *LoG* de 7×7 píxeles. Es decir, que se tienen 49 elementos, cada uno con un valor diferente calculado a partir de la fórmula *LoG* descrita en (9.19). Esto está en relación directa con el tamaño del filtro gaussiano y, obviamente, con las derivadas parciales de segundo orden del mismo.

Aunque **Deng y Pinoli** no establecieron explícitamente un filtro *LIP-LoG*, sin embargo, es fácil extender el mecanismo que los autores utilizaron para definir la versión original del *LIP-Sobel* para obtener un mecanismo análogo y obtener un filtro *LIP-LoG*, cuya fórmula viene dada por:

$$\text{LoG}_{\Delta}^{7 \times 7}(\hat{f}) = \prod_{i=1}^7 \prod_{j=1}^7 \left(H(i, j) \Delta \hat{f}(i, j) \right) \quad (9.21)$$

que se traduce a la siguiente fórmula:

$$\text{LoG}_{\Delta}^{7 \times 7}(\hat{f}) = \prod_{i=1}^7 \prod_{j=1}^7 \left(f(i, j)^{H(i, j)} \right), \quad (9.22)$$

donde $H(i, j)$ es el filtro *LoG* bidimensional definido en (9.16) y $f(i, j)$ son cada uno de los 49 tonos de gris diferentes de la imagen de cada convolución.

Por el mismo motivo que en el caso del método *LIP-Sobel*, se ha comprobado el funcionamiento de estos operadores que se han descrito,



Figura 9.5
Imagen Lenna original.



Figura 9.6
Imagen Lenna oscurecida.

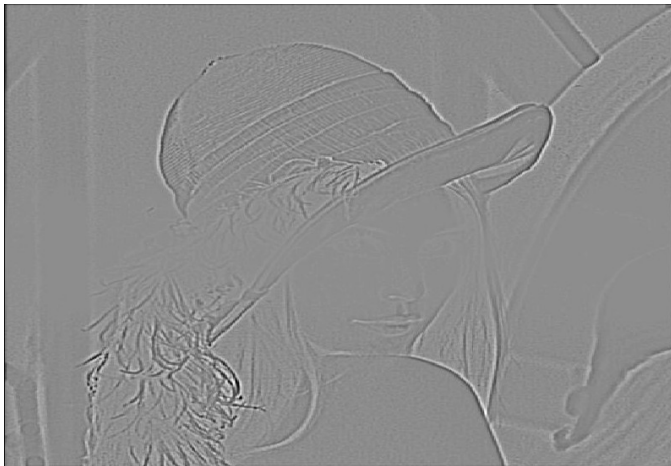


Figura 9.7
LoG sobre la imagen
Lenna oscurecida.

utilizando una imagen conocida y de uso común, Lenna (Fig. 9.5). Esta imagen ha sido modificada mediante la aplicación una banda de oscurecimiento sobre los ojos (Fig. 9.6). Se ha calculado el filtro *LoG* (con $\sigma = 1.0$) aplicando (9.15) sobre la imagen oscurecida, obteniéndose Fig.

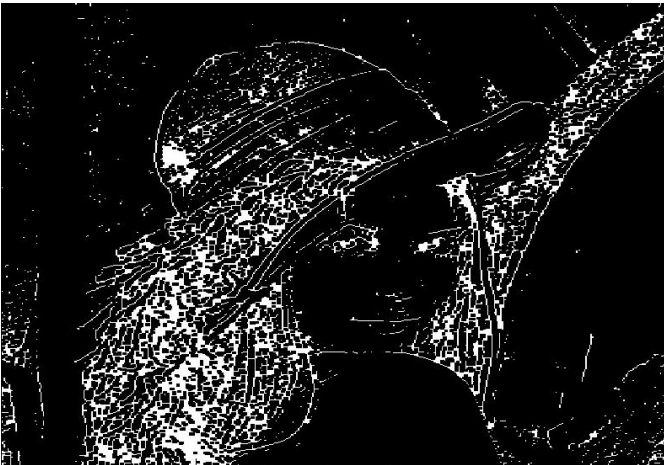
Figura 9.8
LoG umbralizado
(con umbral = 10)
sobre la imagen Lenna
oscurecida.



Figura 9.9
LoG-LIP sobre la ima-
gen Lenna oscurecida.



Figura 9.10
LoG-LIP umbralizado
(con umbral = 10)
sobre la imagen Lenna
oscurecida.



9.7, que ha sido umbralizada con un umbral = 10 y se ha generado la Fig. 9.8. Por otra parte, se ha aplicado el operador *LIP-LoG* (con $\sigma = 1.0$), obteniéndose Fig. 9.9, que ha sido umbralizada también con un valor de umbral = 10 generando la Fig. 9.10. Como se puede ob-



Figura 9.11
Imagen de un so-
portal con sombras
naturales.

servar por simple inspección, la imagen umbralizada tras el operador *LIP-LoG* proporciona bordes en la zona oscurecida, frente a la imagen umbralizada tras aplicar el operador *LoG*, el cual no obtiene bordes en dicha zona.

9.3.2 *LIP-Canny*

El método de *Canny* [17] es un método muy robusto para la extracción de bordes en imágenes naturales que pueden tener presencia de ruido blanco gaussiano. En este apartado se propone un experimento que consiste en la aplicación del paradigma LIP al extractor de bordes de *Canny*. La adaptación a LIP se realiza en los primeros pasos, la obtención de los mapas de gradiente, en lugar de utilizar el operador *convolución* estándar se utiliza el operador *Convolución-LIP* expuesto en (6.13). En el experimento realizado, se ha obtenido el mapa de bordes mediante tres mecanismos:

- Utilizando únicamente el método de *Canny* tradicional.
- Mediante una técnica híbrida en la que primero se aplica *Filtrado Homomórfico* y posteriormente el método de *Canny* tradicional.
- A través de un nuevo método propuesto, *LIP-Canny*.

Debido a la necesidad de comprobar el correcto funcionamiento en imágenes con sombras proyectadas reales, para este experimento se va a utilizar una de las imágenes seleccionadas en el Cap. 7. Se ha escogido la Imagen 4 del DOMINIO exteriores, que se muestra en la Fig. 9.11. Como puede observarse, esta imagen representa una fotografía de un soportal, parte del cual está en sombras. Utilizando esta imagen, se ha obtenido la Fig. 9.12 que muestra los bordes mediante el método tradicional de *Canny*. La Fig. 9.13 contiene los contornos obtenidos mediante la aplicación del *Filtrado Homomórfico* seguido del método de *Canny* habitual y por último, los bordes utilizando el nuevo método propuesto *LIP-Canny* se muestran en la Fig. 9.14. Para todos los métodos se han utilizado los valores de los parámetros que han proporcionado en cada caso los mejores resultados, calculados estos mediante

Figura 9.12
Mapa de bordes
obtenidos utilizando
Canny. Parámetros:
 $th_{low} = 0.05$, $th_{high} =$
 0.125 .

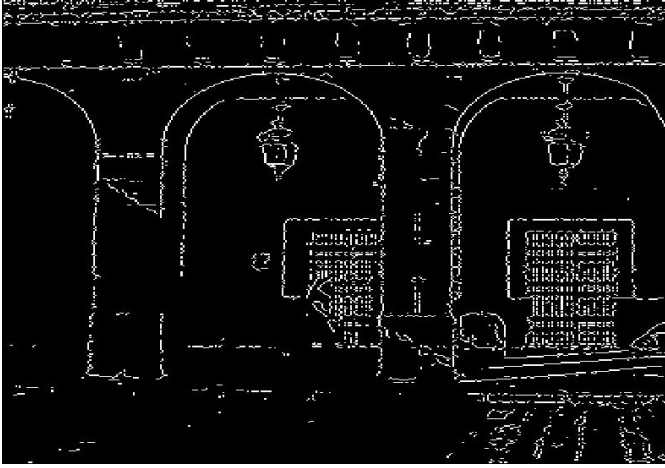


Figura 9.13
Mapa de bordes
obtenidos utilizando
Filtrado Homomórfico
seguido de *Canny*.
Parámetros: $th_{low} =$
 0.01 , $th_{high} = 0.05$.

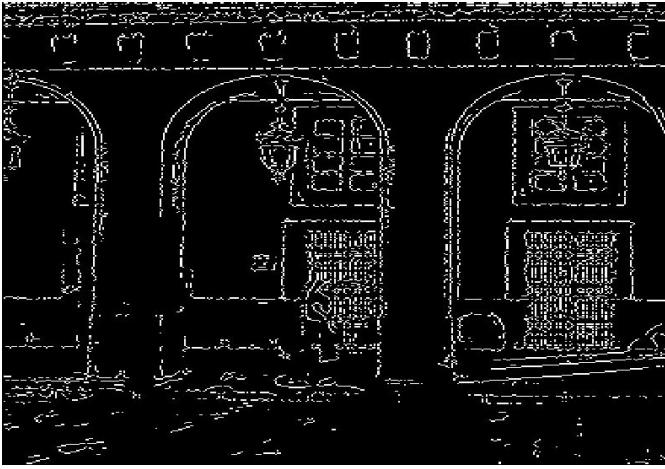


Figura 9.14
Mapa de bordes ob-
tenidos utilizando
LIP-Canny. Paráme-
tros: $th_{low} = 0.05$,
 $th_{high} = 0.4$.



métodos subjetivos. De manera común a todos los métodos, se ha tomado $\sigma = 1.0$, donde σ es el valor de la amplitud de la campana de la gaussiana utilizada en las primeras fases de *convolución*.

Como se puede observar, el método *LIP-Canny* obtiene una mayor cantidad de bordes que el método de *Canny* para el mismo valor de umbral. Es más, el método *LIP-Canny* obtiene bordes precisos en la zona donde se proyecta la sombra, en la cual, el método tradicional de *Canny* no detecta ningún borde. Para que el algoritmo tradicional de *Canny* sea capaz de detectar dichos bordes, es necesario aplicar un paso previo de *Filtrado Homomórfico*. En este caso, el ajuste del umbral tiene que ser mucho más preciso, puesto que los valores del gradiente con los que opera *Canny* tras el *Filtrado Homomórfico* son mucho más pequeños.

En la imagen de contornos obtenidos mediante *LIP-Canny*, Fig. 9.14, se muestran los bordes de las cristaleras de la pared trasera de los soporales, que se encuentran en una zona de sombras en la imagen original, Fig. 9.11. Por el contrario, en la Fig. 9.12, que contiene el mapa de bordes obtenidos por *Canny*, no aparecen esos mismos elementos.

Hay que indicar que los parámetros de los umbrales en el método de *LIP-Canny* son más altos que los mismos para el método de *Canny*, ya que para el mismo nivel de los umbrales, *LIP-Canny* es capaz de detectar muchos más contornos que *Canny*. Esto podría, en algunos casos, llegar a presentar un problema, porque podría detectar como falsos bordes, los píxeles ruidosos de las zonas oscuras. Por otro lado, al aplicar el algoritmo compuesto por el *Filtrado Homomórfico* seguido del método de *Canny* aparece un gran problema de sensibilidad del umbral superior. Es decir, se dejan de detectar la gran mayoría de bordes si se selecciona un umbral algo mayor que el óptimo, y al contrario, se comienzan a detectar una cantidad excesiva de contornos en cuanto se desciende por debajo de dicho umbral. Lo que, en definitiva, hace que *LIP-Canny* se pueda considerar como el mecanismo más robusto y versátil de los 3 comparados en este experimento.

9.4 TIEMPOS DE CÓMPUTO

En las secciones anteriores, se ha mostrado que los algoritmos implementados mediante el operador propuesto *Convolución-LIP* presentan una respuesta visual mejor en todos los casos que los algoritmos equivalentes No LIP. En esta sección, se muestran los tiempos de cómputo obtenidos para los diversos experimentos descritos con anterioridad y la comparación de los diferentes métodos en consideración. En general, los tiempos de cálculo que se muestran en este trabajo se han obtenido implementando todos los algoritmos y métodos utilizando MATLAB[®]7 (R14) utilizando la Image Processing Toolbox y ejecutando los mismos en igualdad de carga computacional sobre un sistema Pentium[®]Centrino M725 a 1.60 GHz (2MB L2-Cache) con 512 MB RAM.

Esta sección describe los resultados obtenidos en cada uno de los experimentos sin entrar en valoraciones o análisis de mayor profundidad, que se realizan en la Sec. 9.5.

9.4.1 Aumento del rendimiento (Speedup)

Además de mostrar el tiempo de cálculo de cada experimento, también se va a incluir otra medida, la ganancia de velocidad o *Speedup*, que se nota como S , y se calcula mediante la siguiente fórmula matemática:

$$S = \frac{\text{Time}_{\text{reference}}}{\text{Time}_{\text{alternative}}} \quad (9.23)$$

Este término describe la velocidad relativa de un sistema nuevo en relación a otro sistema, que es tomado como referencia. Si el valor de ganancia de velocidad es superior a uno, significa que el nuevo sistema es más rápido que el de referencia; y al contrario, si el *Speedup* es inferior a uno, el nuevo sistema se comporta de una manera más lenta que el de referencia.

9.4.2 LIP-Sobel

Se han realizado una serie de experimentos con *LIP-Sobel*, que se han descrito anteriormente. En la Tab. 9.2 se muestra el tiempo de cálculo para el algoritmo *LIP-Sobel*, implementado en las tres versiones en consideración (Deng y Pinoli, DGLIP-CONV y FGLIP-CONV) y con imágenes de dos tamaños diferentes (512×512 y 320×240).

Aunque en la Sec. 9.5.1 se realiza un análisis más profundo de estos datos, es conveniente indicar que para este experimento en particular, la versión más rápida es la implementada utilizando FGLIP-CONV. Por el contrario, la versión de *LIP-Sobel* mediante DGLIP-CONV es más lenta que la implementación propuesta por Deng y Pinoli. Esta situación ocurre debido a que, como se ha indicado en la Sec. 9.1, la formulación obtenida por Deng y Pinoli es una instanciación concreta, optimizada para que dicha formulación sea compacta. A su vez, dicha formulación es equivalente a la obtenida por el mecanismo DGLIP-CONV, que sin embargo, es de carácter general y no se encuentra optimizada para la minimización en el número de operaciones.

Tabla 9.2
Comparación de
tiempos de cálculo
(tiempo en segun-
dos).

ALGORITMO	TAMAÑO	MÉTODO	TIEMPO
LIP-Sobel	512×512	Deng y Pinoli	1.008
		FGLIP-CONV	0.450
		DGLIP-CONV	5.847
	320×240	Deng y Pinoli	0.386
		FGLIP-CONV	0.155
		DGLIP-CONV	1.682

Aumento del rendimiento (Speedup)

En la Tab. 9.3 se muestra el aumento del rendimiento de *LIP-Sobel* implementado con FGLIP-CONV y DGLIP-CONV, en relación con el método original propuesto por Deng y Pinoli.

ALGORITMO	TAMAÑO	SPEEDUP	
		FGLIP-CONV	DGLIP-CONV
LIP-Sobel	512 × 512	2.24	0.17
	320 × 240	2.49	0.23

Tabla 9.3
Speedup del LIP-Sobel con respecto al mismo implementado mediante la técnica de Deng y Pinoli.

9.4.3 Algoritmos de emborronamiento: LIP-Avg, LIP-GBLur

Los algoritmos de emborronamiento se han agrupado y se muestran tanto los tiempos de cálculo como el *Speedup* (este último siempre tomando como referencia el tiempo obtenido por el algoritmo implementado mediante la técnica de Deng y Pinoli). Se han realizado tres experimentos diferentes:

- Emborronamiento por filtro media de tamaño 3×3 píxeles.
- Emborronamiento por filtro media de tamaño 5×5 píxeles.
- Emborronamiento por filtro gaussiano de tamaño 7×7 píxeles.

Estos experimentos se han realizado para dos tamaños de imágenes diferentes (512×512 y 320×240 píxeles). Y se han realizado mediante las tres técnicas LIP en consideración:

- El método propuesto por Deng y Pinoli.
- Mediante *Convolución-LIP* implementado utilizando:
 - DGLIP-CONV.
 - FGLIP-CONV.

Tiempos de Cálculo

En la Tab. 9.4 se muestran los tiempos de los distintos experimentos sobre la misma arquitectura que la descrita para el experimento LIP-Sobel. En esta tabla se vuelve a mostrar, que para todos los algoritmos de este experimento, las versiones más rápidas son aquellas implementadas mediante FGLIP-CONV. Por el contrario, tanto para los algoritmos LIP-Avg (3×3) como para LIP-Avg (5×5), al igual que ocurría en el experimento LIP-Sobel, las implementaciones de los algoritmos que utilizan DGLIP-CONV son las más lentas. Sin embargo, para el algoritmo LIP-GBLur, la implementación mediante el mecanismo propuesto por Deng y Pinoli es claramente la más lenta, ya que no puede aprovecharse de la optimización y compactación de la fórmula matemática que trae dicha propuesta y, por el contrario, tanto FGLIP-CONV como DGLIP-CONV pueden hacer uso de la ventaja computacional que les ofrece la separabilidad de los filtros.

Aumento del rendimiento (Speedup)

De manera similar a lo expuesto en el experimento anterior, en la Tab. 9.5 se indica el *Speedup*, tomando como referencia el tiempo de cada experimento implementado utilizando el método propuesto por Deng y Pinoli.

Tabla 9.4
Comparación de
tiempos de cálculo
(tiempo en segun-
dos).

ALGORITMO	TAMAÑO	MÉTODO	TIEMPO
LIP-Avg (3 × 3)	512 × 512	Deng y Pinoli	0.456
		FGLIP-CONV	0.192
		DGLIP-CONV	3.448
	320 × 240	Deng y Pinoli	0.227
		FGLIP-CONV	0.076
		DGLIP-CONV	1.017
LIP-Avg (5 × 5)	512 × 512	Deng y Pinoli	0.482
		FGLIP-CONV	0.198
		DGLIP-CONV	4.056
	320 × 240	Deng y Pinoli	0.246
		FGLIP-CONV	0.077
		DGLIP-CONV	1.189
LIP-GBlur	512 × 512	Deng y Pinoli	9.629
		FGLIP-CONV	0.333
		DGLIP-CONV	4.560
	320 × 240	Deng y Pinoli	2.855
		FGLIP-CONV	0.175
		DGLIP-CONV	1.421

Tabla 9.5
Speedup para cada
experimento y mé-
todo con respecto al
mismo implementado
mediante la técnica
de Deng y Pinoli.

ALGORITMO	TAMAÑO	SPEEDUP	
		FGLIP-CONV	DGLIP-CONV
LIP-Avg (3 × 3)	512 × 512	2.38	0.13
	320 × 240	2.99	0.22
LIP-Avg (5 × 5)	512 × 512	2.43	0.12
	320 × 240	3.20	0.21
LIP-GBlur	512 × 512	28.92	2.11
	320 × 240	16.31	2.01

9.4.4 Extracción de bordes: LIP-LoG

En este apartado se muestran los resultados obtenidos al implementar el algoritmo *LIP-LoG*. Para este algoritmo se han realizado otros experimentos diferentes a los casos anteriores; se compara el algoritmo original No LIP, es decir, el algoritmo original *LoG*, con los algoritmos LIP tanto por el método de **Deng y Pinoli**, como por *Convolución-LIP* utilizando la aproximación FGLIP-CONV, que se ha demostrado ser la aproximación más rápida. Las imágenes utilizadas tienen tamaños de 512 × 512 y de 256 × 256 píxeles.

Tiempo de Cálculo

En la Tab. 9.6 se muestran los tiempos de cálculo para los diferentes experimentos. Se observa que FGLIP-CONV es la implementación más

TAMAÑO	MÉTODO	TIEMPO
512 × 512	Standard LoG	0.1357
	Deng y Pinoli	7.8312
	FGLIP-CONV	0.2163
256 × 256	Standard LoG	0.0390
	Deng y Pinoli	1.2042
	FGLIP-CONV	0.0636

Tabla 9.6
Comparación de
tiempos de cálculo
(tiempo en segundos)
del experimento LoG
y LIP-LoG.

rápida, mientras que la implementación mediante el mecanismo propuesto por Deng y Pinoli es el más lento, manteniéndose el método DGLIP-CONV entre ambos al comparar sus tiempos de cálculo.

Aumento del rendimiento (Speedup)

Para este experimento se indican dos medidas de *Speedup*. El primero, SPEEDUP D&P, tomando como referencia el método de Deng y Pinoli; mientras que el segundo (denominado SPEEDUP FG.), toma como referencia el método FGLIP-CONV. Los valores se muestran en la Tab. 9.7.

TAMAÑO	MÉTODO	SPEEDUP D&P.	SPEEDUP FG.
512 × 512	Standard LoG	57.71	1.59
	Deng y Pinoli	1.00	0.03
	FGLIP-CONV	36.20	1.00
256 × 256	Standard LoG	30.88	1.63
	Deng y Pinoli	1.00	0.05
	FGLIP-CONV	18.93	1.00

Tabla 9.7
Comparación de
Speedup del expe-
rimento LoG y LIP-
LoG.

9.4.5 Extracción de bordes: LIP-Canny

Este experimento, en el que se utiliza el algoritmo de extracción de bordes de Canny, compara dos aproximaciones invariantes ante cambios de iluminación y la versión original del algoritmo. Las dos aproximaciones invariantes ante cambios de iluminación son:

- El algoritmo LIP-Canny utilizando FGLIP-CONV, tal y como se ha indicado en Sec. 9.3.2.
- Un preprocesamiento previo de la imagen mediante *Filtrado Homomórfico* y posterior aplicación del algoritmo estándar de Canny sobre la imagen reflectancia obtenida del paso de preprocesamiento anterior.

Las imágenes son diferentes de las utilizadas con anterioridad, ya que debían incluir sombras proyectadas naturales. Estas nuevas imágenes tienen tamaños de 425 × 640 y 374 × 353 píxeles, respectivamente.

Tiempo de Cálculo

En la Tab. 9.8 se muestran los tiempos de cálculo para cada uno de los experimentos e imágenes. En este experimento, la versión más rápida es la que aplica *Canny* directamente, aunque como se ha indicado en la Sec. 9.3.2, ésta no es capaz de proporcionar los contornos de las zonas en sombras. Entre las dos implementaciones que permiten la obtención de los bordes de las zonas oscuras, el método LIP-CANNY es más rápido que el mecanismo híbrido constituido por la aplicación de un *Filtrado Homomórfico* seguido del algoritmo de *Canny*, notado como *F.H. + Canny*.

Tabla 9.8
Tiempos de procesamiento para *Canny*, *LIP-Canny* y *Filtrado Homomórfico (F.H.)* seguido de *Canny*.

TAMAÑO	MÉTODO	TIEMPO
425 × 640	Canny	1.190
	F.H. + Canny	1.341
	LIP-CANNY	1.250
374 × 353	Canny	0.519
	F.H. + Canny	0.623
	LIP-CANNY	0.585

Aumento del rendimiento (Speedup)

En la Tab. 9.9 se muestra el *Speedup* de cada uno de los algoritmos, tomando el *LIP-Canny* como sistema de referencia para cada tamaño de imagen.

Tabla 9.9
Speedup para *Canny*, *LIP-Canny* y *Filtrado Homomórfico (F.H.)* seguido de *Canny*.

TAMAÑO	MÉTODO	SPEEDUP
425 × 640	Canny	0.952
	F.H. + Canny	1.073
	LIP-CANNY	1.0
374 × 353	Canny	0.887
	F.H. + Canny	1.065
	LIP-CANNY	1.0

9.5 RESUMEN DE RESULTADOS

Analizando los resultados obtenidos se puede estimar la calidad de los métodos propuestos, al menos, desde un punto de vista de eficiencia computacional y de tiempo de procesamiento. En los siguientes apartados, se van a tratar cada uno de los experimentos, calculando en cada caso el aumento de rendimiento general de cada implementación.

9.5.1 LIP-Sobel

Si se toma únicamente el aspecto visual, se puede observar que el algoritmo estándar de *Sobel* no es capaz de mostrar claramente los contornos de las zonas más oscuras, mientras que la implementación *LIP-Sobel* sí que muestra los bordes de toda la imagen, de manera independiente al grado de iluminación local.

Con respecto al tiempo de procesamiento y al aumento de rendimiento, la ganancia media del experimento implementado mediante FGLIP-CONV es aproximadamente de 2.5 veces más que el mismo implementado por el método de **Deng y Pinoli**. Mientras que, en cambio, el *LIP-Sobel* utilizando el método DGLIP-CONV es más lento que el mismo siguiendo la propuesta de **Deng y Pinoli**, ya que se obtiene un *Speedup* de 0.21, es decir, es aproximadamente un 80 % más lento.

Este experimento requiere un análisis más profundo de los resultados temporales, ya que a pesar de que la versión DGLIP-CONV del algoritmo *LIP-Sobel* utiliza un mecanismo de *convolución* por filtros separables, es más lenta que la implementada mediante el mecanismo propuesto por **Deng y Pinoli**, que utiliza filtros bidimensionales, que haría que su aplicación fuese teóricamente más lenta. La lentitud de la propuesta DGLIP-CONV viene dada por dos factores: la elevada carga computacional de los operadores multiplicación y potenciación, y el hecho de que el filtro bidimensional de Sobel contiene muchos valores a 0. Esta segunda característica hace que, para cada píxel, la aplicación del filtro bidimensional de *LIP-Sobel* mediante el mecanismo propuesto por **Deng y Pinoli** se realice con seis multiplicaciones, una división y dos potenciaciones al cuadrado (que podrían implementarse mediante dos multiplicaciones o de manera aún más efectiva mediante un desplazamiento binario de un bit hacia la izquierda). Por tanto, tomando de manera simplificada que el tamaño de la imagen es de $n \times n$ píxeles, la aplicación de *LIP-Sobel* según el método propuesto por **Deng y Pinoli** proporciona 18 operaciones por píxel, como muestra (9.24). En cualquier caso, se necesita el doble de las operaciones descritas ya que para obtener *LIP-Sobel* hay que aplicar dos filtros bidimensionales para el cálculo de la imagen gradiente. Para poder proporcionar datos aún más precisos, se ha podido estimar que el operador potenciación es cinco veces más lento que el operador multiplicación, por tanto, para poder realizar comparaciones, se va a calcular el número de *operaciones equivalentes* (oper. eq.) de cada implementación. La normalización en *operaciones equivalentes* hace que una potenciación sea igual a 5 oper. eq. y que 1 multiplicación o división coincida con 1 oper. eq.

$$\begin{aligned} \text{Número de operaciones de LIP - Sobel}_{\text{Deng\&Pinoli}} : \\ 2 \times (n^2 \times (6 \text{ multip.} + 1 \text{ divis.} + 2 \text{ potenc.})) \Rightarrow \\ 2 \times n^2 \times 9 \text{ oper.} \Rightarrow 18n^2 \text{ oper.} \Rightarrow 34n^2 \text{ oper. eq.} \end{aligned} \quad (9.24)$$

Sin embargo, la implementación de *LIP-Sobel* utilizando DGLIP-CONV requiere, para una imagen de igual tamaño, una doble aplicación de 3 multiplicaciones y 3 potenciaciones por filtro. En (9.25) se muestra la estimación del número de operaciones requeridas para la implementación del que el número de operaciones por píxel es menor que en el

caso anterior, pero el número total para la imagen completa es mayor. También se indica el número de *operaciones equivalentes*.

$$\begin{aligned} & \text{Número de operaciones de LIP - Sobel}_{\text{DGLIP-Conv}} : \\ & 2 \times (2 \times (n^2 \times (3 \text{ multíp.} + 3 \text{ potenc.}))) \Rightarrow \\ & 4 \times n^2 \times 6 \text{ oper.} \Rightarrow 24n^2 \text{ oper.} \Rightarrow 72n^2 \text{ oper. eq.} \end{aligned} \quad (9.25)$$

Para DGLIP-CONV, el cálculo utiliza un filtro en el que cada elemento del mismo tiene un valor diferente no nulo, por lo cual, no es posible realizar ninguna optimización para obtener una reducción en el número de potenciaciones (que es la operación computacionalmente más costosa). Como se podrá comprobar en otros experimentos, en cuanto los filtros que se aplican tienen un mayor tamaño y están compuestos por valores no uniformes y no nulos, la ventaja de la propuesta de **Deng y Pinoli** se esfuma y la implementación DGLIP-CONV también muestra una mejora en los tiempos de ejecución frente a dicha propuesta.

9.5.2 LIP-Avg, LIP-GBLur

Como se ha comentado a lo largo del capítulo, estos dos experimentos muestran los dos extremos de *caso más favorable* y *más desfavorable* para el método propuesto por **Deng y Pinoli**. El algoritmo *LIP-Avg* realiza el filtro *media* y se corresponde con el “mejor caso”, ya que se puede obtener un exponente común para todos los píxeles. Frente a este caso se encuentra el del algoritmo *LIP-GBLur*, que aplica un filtrado gaussiano y es el “peor caso”, ya que no es posible sacar un exponente común y hay que elevar cada píxel a un exponente diferente. Si se calcula el *Speedup* de cada experimento, según tamaño de imagen y de filtro, de la implementación utilizando el método propuesto por **Deng y Pinoli** frente a las implementaciones de los mismos utilizando DGLIP-CONV y FGLIP-CONV, se puede observar que para el experimento *LIP-Avg* es de 2.75 y de 0.17, para FGLIP-CONV y para DGLIP-CONV, respectivamente. Por otra parte, para el experimento *LIP-GBLur* se obtiene un aumento del rendimiento de 22.62 para FGLIP-CONV y de 2.06 para DGLIP-CONV. Así, en general, el *Speedup* medio para los dos experimentos (para todos los tamaños de imagen y de filtros) del método FGLIP-CONV es de 9.38; mientras que el *Speedup* medio de los experimentos implementados utilizando DGLIP-CONV alcanza un modesto 0.80. Esto demuestra que la implementación mediante FGLIP-CONV es más rápida que el mecanismo propuesto por **Deng y Pinoli**, mientras que la implementación mediante DGLIP-CONV se comporta de forma más lenta.

En el experimento *LIP-Avg* (3×3), el denominado “mejor caso” para el mecanismo propuesto por **Deng y Pinoli**, se aplican 9 multiplicaciones y una potenciación (10 operaciones) por cada píxel, por lo tanto, esta implementación requiere 14 *operaciones equivalentes*.

$$\begin{aligned} & \text{Número de operaciones de LIP - Avg}(3 \times 3)_{\text{Deng\&Pinoli}} : \\ & n^2 \times (9 \text{ multíp.} + 1 \text{ potenc.}) \Rightarrow 10n^2 \text{ oper.} \Rightarrow 14n^2 \text{ oper. eq.} \end{aligned} \quad (9.26)$$

Por otra parte, en el método DGLIP-CONV se necesitan 3 multiplicaciones y 3 potenciaciones, pero hay que aplicar dos veces dicho filtrado (12 operaciones) en cada píxel, que implican 36 *operaciones equivalentes*. En (9.26) y (9.27), se muestran respectivamente los cálculos del número de operaciones aplicadas en ambos métodos.

$$\begin{aligned} & \text{Número de operaciones de LIP - Avg}(3 \times 3)_{\text{DGLIP-Conv}} : \\ & 2 \times (n^2 \times (3 \text{ multip.} + 3 \text{ potenc.})) \Rightarrow \\ & 2 \times n^2 \times 6 \text{ oper.} \Rightarrow 12n^2 \text{ oper.} \Rightarrow 36n^2 \text{ oper. eq.} \end{aligned} \quad (9.27)$$

Por el contrario, en el último experimento LIP-GBLUR, la implementación mediante el mecanismo propuesto por Deng y Pinoli utiliza 49 multiplicaciones y 49 potenciaciones (98 operaciones) para calcular cada nuevo píxel, tal y como se muestra en (9.28), la cantidad de operaciones totales es de $294n^2$ *operaciones equivalentes*.

$$\begin{aligned} & \text{Número de operaciones de LIP - GBLUR}_{\text{Deng\&Pinoli}} : \\ & n^2 \times (49 \text{ multip.} + 49 \text{ potenc.}) \Rightarrow 98n^2 \text{ oper.} \Rightarrow 294n^2 \text{ oper. eq.} \end{aligned} \quad (9.28)$$

Mientras que para la versión que utiliza DGLIP-CONV, la cantidad de operaciones se reduce a 7 multiplicaciones y 7 potenciaciones por píxel y filtro. En (9.29) se muestra que el número de *operaciones equivalentes* total es de $84n^2$.

$$\begin{aligned} & \text{Número de operaciones de LIP - GBLUR}_{\text{DGLIP-Conv}} : \\ & 2 \times (n^2 \times (7 \text{ multip.} + 7 \text{ potenc.})) \Rightarrow \\ & 2 \times n^2 \times 14 \text{ oper.} \Rightarrow 28n^2 \text{ oper.} \Rightarrow 84n^2 \text{ oper. eq.} \end{aligned} \quad (9.29)$$

En este caso, el hecho de que DGLIP-CONV haga uso de un menor número de operaciones unido al uso de los filtros separables, tiene mayor importancia. Gracias a una mejor política de acceso (debido a un menor número de fallos de caché), esta implementación consigue batir holgadamente al método propuesto por Deng y Pinoli.

9.5.3 LIP-LoG

Tras observar en los experimentos anteriores que FGLIP-CONV es el método más rápido para la implementación del operador *Convolución-LIP*, para este algoritmo (*LIP-LoG*) sólo se han realizado pruebas mediante el algoritmo *LoG* estándar (no LIP), *LIP-LoG* implementado mediante la técnica adaptada de la propuesta de Deng y Pinoli y *LIP-LoG* implementado utilizando FGLIP-CONV. En general, se puede observar que *LIP-LoG* es capaz de detectar bordes en zonas oscurecidas, mientras que el algoritmo *LoG* estándar no es capaz de encontrar dichos contornos. Por lo que, en el aspecto temporal, aunque *LoG* es el algoritmo que menos tarda, no proporciona una respuesta satisfactoria, al perderse contornos de zonas oscuras, que sí son detectados por *LIP-LoG*. Particularizando en las dos implementaciones del *LIP-LoG*, que son mediante el método de Deng y Pinoli y mediante FGLIP-CONV,

se obtienen las siguientes medidas de ganancia de velocidad: FGLIP-CONV es aproximadamente un 60 % más lento que el algoritmo básico de LoG, mientras que es entre un 200 % y un 350 % más rápido que el método propuesto por Deng y Pinoli.

9.5.4 LIP-Canny

Como ya se ha comentado en la Sec. 9.3.2, el algoritmo de Canny no es capaz de detectar contornos en las zonas oscuras ni en las sombras proyectadas, sin embargo, la implementación de LIP-Canny realizada mediante FGLIP-CONV obtiene todos los contornos de los elementos de dichas zonas. Por tanto, se puede afirmar que LIP-Canny mejora el funcionamiento general del método de Canny.

Con respecto a los tiempos de cómputo, se han realizado experimentos para dos tamaños distintos (la de tamaño menor, es una imagen recortada de la mayor). El tiempo de cálculo de Canny es aproximadamente entre un 14 % y un 5 % más rápido que el de LIP-Canny y esto es debido a que para calcular LIP-Canny se realizan una serie de operaciones adicionales, frente a la convolución simple que se realiza en el método de Canny. Para obtener una solución similar, a la cual no le afecten los cambios de iluminación, como se ha visto anteriormente, habría que aplicar un Filtrado Homomórfico. Esto aumenta el tiempo de cómputo con respecto a la aplicación directa del método de Canny. Por los experimentos realizados, se ha obtenido que la aplicación de un Filtrado Homomórfico seguido del método de Canny es entre un 13 % y 20 % más lento que aplicar únicamente el método de Canny.

Para este experimento, se ha calculado el Speedup tomando como tiempo de referencia, según el tamaño de la imagen, el resultado de tiempos de la aplicación del método propuesto de LIP-Canny. Se puede observar, que el método híbrido (primero, Filtrado Homomórfico y posteriormente, el método de Canny) es entre un 6.5 y un 7.3 % más lento que LIP-Canny. Por lo tanto, también se puede afirmar que ambas implementaciones proporcionan resultados equiparables entre sí, aunque el mecanismo híbrido es más lento y el ajuste del umbral de los parámetros para dicho método es mucho más sensible a pequeñas modificaciones que en el caso de LIP-Canny, que se comporta de forma más robusta.

9.6 CONCLUSIONES

En los apartados anteriores se ha mostrado el nuevo operador propuesto (Convolución-LIP), que puede ser implementado mediante dos versiones diferentes (DGLIP-CONV y FGLIP-CONV). Con ellos se ha podido observar que se puede generalizar de manera simple y efectiva cualquier algoritmo de procesamiento de imágenes al paradigma LIP siempre que se utilicen filtros de convolución separables. Aunque en trabajos previos [29], otros autores habían dejado entrever un posible método para este fin, el operador que se propone en esta Tesis Doctoral es más versátil, puesto que permite adaptar cualquier filtro sin necesidad de recodificación.

Además, se han mostrado los tiempos de cómputo de los distintos ex-

perimentos. En ellos, se ha demostrado que las implementaciones de los algoritmos realizadas mediante FGLIP-CONV son las más rápidas, lográndose aumentos de velocidad (*Speedup*) medias de 200 %, es decir, el doble de rápido que otras implementaciones. En algunos experimentos se han obtenido aumentos de velocidad de más de 3600 %.

RESULTADOS DE LOS CUESTIONARIOS DE CALIDAD DE BORDES

ÍNDICE

10.1	Cuestionarios para evaluación de contornos	184
10.2	Pilotaje	185
10.2.1	Pilotaje previo	185
10.2.2	Pilotaje final	186
10.3	Visibilidad de la encuesta	187
10.4	Encuestados	188
10.4.1	Muestra poblacional	188
10.4.2	Carta de presentación	189
10.4.3	Tasa de respuesta	192
10.5	Resultados del proceso de evaluación	194
10.5.1	Respuestas cualitativas	194
10.5.2	Metodología de las respuestas	195
10.6	Análisis preliminar de los cuestionarios	196
10.6.1	Análisis de la población	196
10.6.2	Muestreo a posteriori	197
10.6.3	Análisis de frecuencia de cada pregunta	198
10.7	Análisis de los resultados	200
10.7.1	Análisis por grado de visión de los encuestados	200
10.8	Validación de la encuesta	205
10.8.1	Validez de criterio	206
10.8.2	Validez de constructo	206
10.8.3	Validez de contenido	207
10.9	Fiabilidad de la encuesta	208
10.9.1	Fiabilidad externa	208
10.9.2	Fiabilidad interna	208
10.10	Conclusiones	211
10.10.1	Conclusiones sobre el proceso de respuesta	212
10.10.2	Conclusiones sobre los resultados	214
10.10.3	Conclusiones de la Parte III	216

Una de las facetas más singulares del ser humano, que ha marcado su evolución, ha sido la búsqueda del óptimo en todos los ámbitos de interés. Cuando no se puede obtener dicho óptimo, se buscan los métodos que proporcionen los mejores resultados relativos. Sin embargo, cuando la determinación del mejor caso no es fácil de obtener o cuando depende de factores subjetivos no mesurables, el mejor sistema para obtener la calificación es mediante cuestionarios de evaluación. Un claro ejemplo de la utilidad de este sistema se expone en este capítulo, en el que se utilizan dos encuestas para evaluar la calidad de 6 métodos de extracción de bordes diferentes con 2 parámetros de funcionamiento diferentes.

10.1 CUESTIONARIOS PARA EVALUACIÓN DE CONTORNOS

Como se ha comentado a lo largo de esta Tesis, la calidad de los contornos es un concepto con un alto grado de subjetividad: algunos usuarios pueden preferir contornos suaves aunque menos precisos y otros pueden preferir contornos con muy alta precisión aunque sean más abruptos. Es más, la subjetividad no sólo depende de la persona sino también de la utilidad o tarea visual que se desee obtener a partir de dichos contornos: una misma persona puede escoger un mapa de contornos para una determinada tarea y para otra tarea distinta escoger otro mapa de contornos diferente. Por todo esto, se ha determinado que la mejor opción para identificar el algoritmo de extracción de bordes que proporcione el conjunto de contornos globalmente más aceptado es mediante cuestionarios de opinión.

La construcción de una herramienta única para la investigación mediante encuestas de imágenes de contornos es un objetivo que excede el ámbito de esta Tesis. Además, proporcionar una encuesta capaz de evaluar todo tipo de métodos de contornos, de cualquier dominio de imágenes y para cualquier población es poco realista. Más aún cuando ni existen metodologías de diseño de encuestas en este sentido, ni se han desarrollado encuestas previas con las que comparar. Esta Tesis pretende abrir un camino robusto y con indicios de veracidad científica en el diseño y utilización de encuestas para el análisis de la calidad de los métodos de extracción de contornos en imágenes. En este capítulo se muestran los resultados obtenidos a partir de la evaluación subjetiva de la calidad percibida de diferentes algoritmos de extracción de contornos. Sin embargo, en este capítulo, no sólo se proporcionan resultados estrictamente numéricos, sino que también se aportan como resultados de la investigación, toda la información del proceso completo de diseño de los cuestionarios, de la puesta en marcha y de la fase de recolección de respuestas utilizando el sistema Web descrito en el Cap. 8. Dentro del ámbito de la evaluación subjetiva de imágenes y de vídeo, existen procedimientos regulados [30, 54, 55] tanto para construir las encuestas, como para estructurar los mecanismos para contestarlas. Sin embargo, hasta dónde se conoce, con anterioridad a esta Tesis Doctoral la gran mayoría de estudios de investigación mediante cuestionarios que se han realizado para la evaluación de algoritmos de calidad de bordes [31, 47, 50] no han seguido ninguna metodología para la creación de las encuestas, aunque sí utilizan o proponen mecanismos normativos para otras fases de sus respectivos estudios. Por ello, la metodología propuesta en esta Tesis Doctoral para el desarrollo de los cuestionarios es una innovación científica relevante y se incluye en este capítulo como parte del resultado de la investigación.

Todos los criterios de evaluación, que han sido fijados en el Cap. 7, unidos a las recomendaciones para la construcción de cuestionarios que involucren la evaluación de elementos visuales, expuestas en el Cap. 8, se han plasmado en dos encuestas con 22 preguntas cada una. Ambas encuestas son similares entre sí, ya que el texto de las preguntas es el mismo para cada encuesta, variando únicamente las imágenes de contornos que se muestran a los encuestados para su evaluación. Dentro de cada encuesta se hace uso de los distintos tipos de preguntas

descritos en el Cap. 8: preguntas de puntuación con repetición y sin repetición, preguntas de selección con repetición y preguntas con mapeado espacial y temporal de la respuesta. La encuesta contiene tanto preguntas en las que se muestra la imagen original, como preguntas en las que no se muestra la imagen base.

10.2 PILOTAJE

Antes de presentar las dos encuestas definitivas ante los usuarios finales para recabar sus respuestas, se ha producido un proceso largo en el que se han comprobado todos los aspectos y variables estadísticas involucradas. Esta fase, denominada de *pilotaje*, ha sido la base de la mayoría de recomendaciones expuestas en el Cap. 8. El resto de recomendaciones de dicho capítulo provienen del proceso en sí de respuesta de los encuestados con los cuestionarios definitivos. Durante el proceso de pilotaje se han realizado encuestas para encontrar errores, probar y determinar los mejores mecanismos para cada tipo de pregunta, evaluar el grado de entendimiento de las preguntas, etc. Las preguntas utilizadas en estas encuestas se han descartado para las encuestas finales, puesto que estas encuestas de pilotaje no han sido diseñadas con los mismos objetivos que las encuestas finales y, por tanto, los resultados obtenidos no serían coherentes.

Al ser una investigación en un campo nuevo, en el que no se tienen referencias ni metodologías probadas, no es posible diseñar directamente encuestas. En este ámbito, antes de evaluar y ajustar las encuestas definitivas, se ha de obtener alguna metodología, basada en opiniones contrastadas, que permita el diseño de dichas encuestas. Esto ha llevado a que, en esta investigación, el periodo de pilotaje se ha dividido en dos etapas, pilotaje previo y pilotaje final.

10.2.1 *Pilotaje previo*

En la fase de pilotaje previo se han realizado una serie de encuestas en las que se han incluido preguntas con diferentes tipologías. Con estas pruebas de pilotaje se han comprobado, entre otros, la influencia de los diferentes tipos de preguntas en las respuestas de los encuestados, el grado de aceptación del número de imágenes de respuesta por parte de los usuarios o la replicabilidad de las respuestas en preguntas repetidas dentro de una misma encuesta. Estas pruebas de pilotaje no se diseñaron para obtener resultados numéricos cuantificables, sino que se utilizaron para extraer información cualitativa de los encuestados, que se consiguió mediante entrevistas personales. Estas entrevistas personales fueron realizadas de manera individualizada tras cada prueba de pilotaje y siguieron una estructura de preguntas concretas con respuestas abiertas.

Por ejemplo, para determinar el número máximo de imágenes de respuesta por pregunta se construyó una encuesta piloto en la que se incluyeron preguntas con diferentes números de imágenes de respuesta, desde 3 hasta 24 imágenes de respuesta. Se escogieron estos límites puesto que se pretendía evaluar 6 algoritmos diferentes con 4 niveles

de parámetros, aunque posteriormente se redujo a 6 algoritmos y 2 niveles de parámetros. El orden de las preguntas se determinó al azar, cambiándose para cada usuario. Se seleccionaron 5 usuarios para responder a esta encuesta, todos ellos con un grado de visión estándar, con diferentes edades (23, 29, 34, 42, 61 años), y con conocimientos de informática a nivel de usuario medio (todos sabían manejar teclado, ratón, pantalla, páginas Web, etc.). Tras responder a esta encuesta piloto, cuyas respuestas no fueron incluidas en el cálculo del resultado final, se realizó una entrevista personal a cada encuestado. Entre las preguntas formuladas se incluyó:

En función del número de imágenes de respuesta en cada pregunta, ¿cuál de las preguntas le resultó más sencilla de responder?

Se les permitió a todos los usuarios volver a ver las preguntas y todos ellos respondieron que la más complicada era la que presentaba 24 imágenes. Algunos comentarios fueron *“Es inmanejable con tantas imágenes”* o *“Con tantas imágenes se me olvida lo que había visto en las primeras”*. Las preguntas con 3–4 imágenes de respuesta obtuvieron un alto grado de aceptación, aunque el máximo consenso se encontró para 6 imágenes de respuesta. Según los encuestados, con un número bajo de imágenes de respuesta *“tengo pocas imágenes entre las que comparar”*. Por tanto se determinó que, para esta encuesta, en la que se evalúan 6 métodos diferentes, la mejor configuración de preguntas es utilizando un máximo de 6 imágenes de respuesta para las preguntas de puntuación o de selección.

10.2.2 Pilotaje final

Las encuestas fueron testeadas en una prueba de pilotaje por 4 encuestados, tomados al azar entre un grupo amplio de usuarios voluntarios, con diferentes grados de visión y conocimientos informáticos, aunque ninguno presentaba problemas de BAJA VISIÓN. Ninguno de estos 4 voluntarios participaron en la fase de pilotaje previo y sus respuestas fueron descartadas para la fase de encuesta final. Gracias a esta prueba de pilotaje de los cuestionarios se pudieron detectar y eliminar errores, comprobar la comprensión de las preguntas por parte de los encuestados, estimar la coherencia de las respuestas y mejorar el sistema Web de soporte para las encuestas. El pilotaje final se realizó de manera controlada, es decir, anotando los comentarios y sugerencias de los voluntarios. Este control permitió, a su vez, por ejemplo, observar el mecanismo de respuesta, cuánto tiempo tardaron, el proceso de toma de decisiones y las incertidumbres que mostraban ante ciertas preguntas. Los encuestados informaron de los errores que se encontraron. Esto, unido a la evaluación del propio proceso de respuesta de las encuestas que se realizó, permitió un ajuste más fino de las encuestas. Sin embargo, todo esto no garantiza la eliminación completa de todos los errores, ya que en todo caso, sólo se han tenido en cuenta las respuestas de esos encuestados voluntarios, que pueden no haber detectado

todos los problemas existentes. Adicionalmente, puede que algunas de las cuestiones no hayan sido catalogadas como problemáticas por estos encuestados, pero que para otros usuarios pudieran serlas.

10.3 VISIBILIDAD DE LA ENCUESTA

La visibilidad de la encuesta se debe enfocar desde dos ámbitos: facilidad de acceso a las encuestas para los usuarios y tiempo disponible para contestar a la misma. En este estudio, que utiliza un sistema Web para contestar a la encuesta, se ha escogido un servidor Web con los requisitos hardware y software que impone el sistema Web en cuestión y se le ha dotado de una entrada DNS gratuita, elegida por su simplicidad. Al no tener constancia de que el uso de dominios de carácter académico o profesional, pudiera aumentar la visibilidad o el número de respuestas, se entiende que la opción de DNS gratuito es una buena elección.

Al ser relativamente nuevas las encuestas a través de Internet, no hay muchos estudios que indiquen el tiempo que deben permanecer activas las encuestas para que sean contestadas. Los estudios que se realizan mediante encuestas por correo ordinario tienen un tiempo de respuesta relativamente alto (entre 8 y 12 semanas) [115], por lo que, como mínimo hay que mantener durante ese tiempo accesibles los cuestionarios. Este periodo se debe poder modificar a lo largo del mismo en función del número de respuestas, de envíos de solicitudes a nuevos grupos de población o por cualquier imprevisto no contemplado.

En esta investigación en particular, se ha tenido abierta la encuesta en dos periodos. En el primero desde mediados de noviembre hasta mediados de diciembre se realizó el pilotaje final, tras lo que se realizaron las modificaciones pertinentes. Posteriormente, las encuestas se mantuvieron accesibles durante un segundo periodo desde mediados de diciembre hasta primeros de abril. La fecha de cierre inicial de la encuesta, siguiendo la sugerencia indicada anteriormente de 8–12 semanas, se determinó que fuese final de febrero. Sin embargo, debido a que la **TASA DE RESPUESTA** había sido muy baja a falta de 15 días para el cierre de la encuesta, se decidió ampliarla durante un mes más, para cerrarla definitivamente a primeros de abril.

En el siguiente epígrafe se volverá a este tema ya que, además de ampliar el periodo de apertura de la encuesta durante más tiempo, también se decidió utilizar otras técnicas para intentar aumentar la tasa de respuesta, principalmente mediante el envío de recordatorios. En resumen, la encuesta final se ha mantenido abierta y accesible a través de Internet durante 16 semanas (4 meses).

TASA DE RESPUESTA:
Porcentaje de individuos que contestan de manera efectiva una encuesta frente al número total de individuos invitados a contestarla.

10.4 ENCUESTADOS

El universo de encuestados es el conjunto de individuos susceptibles de responder a una determinada encuesta. Cuando una encuesta se aplica a TODA LA POBLACIÓN y se espera respuesta de todos los individuos de la misma, se dice que es un estudio de investigación por *censo*. De manera general no es factible realizar un estudio por *censo* salvo en poblaciones pequeñas y muy controladas, por lo que, lo más habitual es delimitar un subconjunto de individuos que sean representativos de toda la población, que es denominado *muestra poblacional*. Como se ha indicado en el Cap. 4, la muestra poblacional es un aspecto relevante para una investigación fiable basada en encuestas de opinión. La elección de dicha muestra poblacional se puede realizar escogiendo el conjunto de individuos a los que se les realizará la encuesta con carácter previo a que se contesten dichos cuestionarios (*muestreo a priori*), o bien, catalogar cada individuo que haya contestado a la encuesta, aceptando o descartando sus respuestas después de que éstas se hayan realizado (*muestreo a posteriori*). El primer tipo de muestreo se puede realizar cuando se tiene un universo muy amplio y variado de posibles individuos encuestables, y con garantía de que la tasa de respuesta sea alta. Sin embargo, cuando se desconoce el universo de los individuos o cuando se presume que la tasa de respuesta va a ser baja, es preferible realizar un muestreo *a posteriori*, ya que no se puede renunciar a ningún individuo que tenga interés en contestar a las encuestas. Para esta investigación se ha optado por este segundo tipo de muestreo, *a posteriori*, ya que no se conoce *a priori* la tasa de respuesta que puede tener la encuesta, aunque se presupone que podría ser baja.

10.4.1 Muestra poblacional

Una de las principales fuentes de información para el desarrollo de estudios de investigación basados en cuestionarios es el análisis de estudios previos en el mismo ámbito de aplicación o en ámbitos cercanos al que se encuentra bajo estudio. En este sentido, Heath *et al.* [47] es una buena fuente, ya que estos investigadores realizaron un estudio basado en encuestas para determinar qué algoritmo de extracción de contornos y con qué parámetros de funcionamiento proporcionaba los mejores resultados subjetivos. En este sentido, Heath *et al.* escogen un total de 16 participantes para la evaluación final de 28 imágenes originales.

En la población base se pueden identificar dos grupos o *estratos* principales: individuos con VISIÓN ESTÁNDAR e individuos con BAJA VISIÓN. Para cualquiera de los dos grupos, la muestra poblacional podrá incluir individuos de cualquier nivel social, sexual, cultural o de edad, ya que en la hipótesis de trabajo no se consideran estas variables. Por tanto, estos factores no se han tenido en cuenta para la determinación de la muestra poblacional, aunque se ha intentado que la muestra sea lo más amplia y representativa posible. Por motivos de utilidad funcional y de facilidad de selección de la muestra, se han determinado una serie de condicionantes a los factores anteriores. Estos condicionantes han sido escogido para que, *a priori*, no afecten de manera apreciable

a los resultados de la muestra poblacional. En cualquier caso, estos factores no deben ser considerados, en ningún caso, como limitaciones impuestas por motivos científicos, sino por motivos de tipo funcional. Las principales limitaciones han sido:

- Nivel Social: Cualquiera con acceso a Internet, directa o indirectamente.
- Nivel Cultural: Mínimo, saber leer y escribir (equivalente a Educación General Básica, Educación Secundaria o similar).
- Nivel Informático: Saber manejar teclado, ratón y navegar por Internet. En caso negativo, se requiere que algún individuo con dichos conocimientos informáticos actúe como asistente.

La introducción de estos factores limitantes en la selección de la muestra poblacional introduce a su vez, ciertos sesgos que han de ser tenidos en cuenta. El principal es que, de facto, requiere un cierto nivel social y cultural, ya que es imprescindible un ordenador con acceso a Internet. Esto podría afectar al resultado final, aunque se estima que, de manera práctica, no es un factor que desvirtúe excesivamente esta investigación.

Una vez se han especificado las limitaciones para escoger la muestra poblacional, se ha de escoger el mecanismo para acceder a los individuos. Debido a la dificultad en el acceso al estrato de personas con BAJA VISIÓN, además de incluir comunicaciones personales directas con individuos que tengan dicha problemática, se decidió que la mejor solución es a través de organizaciones que traten con ellos, en particular, mediante la ONCE.

Para el estrato de individuos con VISIÓN ESTÁNDAR se ha estimado que el mecanismo más barato y sencillo es mediante la técnica *bola de nieve*, en la que se envían invitaciones a contestar la encuesta a individuos escogidos que tengan acceso a un grupo importante de personas. Estas solicitudes se han realizado mediante cartas de invitación, que en este caso particular se han enviado por email, debido a su reducido coste, su rapidez y su amplia difusión. En los emails se han incluido instrucciones para promover su difusión en cadena (mecanismo de *bola de nieve*), haciendo especial hincapié en la búsqueda de pacientes con BAJA VISIÓN. El principal hándicap que presenta esta técnica es la dificultad que presenta el seguimiento de la tasa de respuestas, por lo que, a lo sumo sólo se puede realizar una estimación aproximativa.

ONCE:
ORGANIZACIÓN
NACIONAL DE CIEGOS
ESPAÑOLES

10.4.2 Carta de presentación

En toda encuesta debe existir una carta en la que se presente la investigación, se explique el método y el funcionamiento del sistema. Según Walonick [115], esta carta es una parte primordial de la investigación, ya que tiene una gran influencia en la decisión de si el individuo responderá o no a la encuesta y también determinará en gran medida la probabilidad de que dicho individuo finalice la misma. En esta investigación se ha escogido la opción de hacer una única carta de presentación que sea válida tanto para individuos con VISIÓN ESTÁNDAR como

para individuos con BAJA VISIÓN. Se ha realizado otra carta de presentación para su envío a la ONCE, en la que se hacía especial énfasis en la BAJA VISIÓN y en cómo esta investigación podría ayudar a los sujetos que la padezcan. En la carta de presentación general (la enviada a la población con VISIÓN ESTÁNDAR y a la población con BAJA VISIÓN) se incluye un sencillo manual de utilización del sistema Web, que describe paso a paso cómo registrarse en el sistema, acceder a las encuestas y responder a las preguntas.

Población inicial

Al usar la técnica de muestreo de *bola de nieve* hay que ser cuidadoso en la elección de la población inicial a la que se le envía la carta de presentación de la encuesta. Se espera que se inicie un reenvío masivo de dicha carta y por tanto, conviene que la población inicial sean individuos que tengan muchas personas a su cargo o con muchos contactos, para que la difusión de la carta de presentación sea amplia y rápida. Esto no siempre se consigue, ya que influyen muchos factores en los individuos de esta población inicial que “cortan la avalancha”: filtrado de intereses, periodos temporalmente muy ocupados, olvido, etc.

La población inicial constaba de 26 individuos, incluyendo a una sección provincial de la ONCE. La población final de usuarios registrados en la Web ha sido de 70 individuos, por lo que el efecto **TASA DE MODIFICACIÓN DE LA POBLACIÓN MUESTRAL** de la técnica *bola de nieve* ha significado un aumento de la población final de un 269,2 %.

Recordatorios

Uno de los principales problemas que presentan las encuestas “online” es que, al no tener un soporte físico, tienden a desaparecer de las prioridades de los individuos, por lo que es conveniente recordarles que las rellenen completamente. Se ha estudiado [115] la influencia de diferentes mecanismos para recordar a los encuestados la necesidad de que contesten a la encuesta, partiendo de que entre un 10 y un 60 % de los individuos suelen responder a las encuestas sin necesidad de recordatorios. Existen diferentes mecanismos de recordatorio, que pueden ir desde el envío de copias de los cuestionarios, llamadas telefónicas personales, reuniones periódicas (en casos de grupos controlados) hasta el envío de postales o emails, entre otros. Una gran mayoría de estudios revelan que el envío de postales o emails de recordatorio suelen aumentar el número de respuestas en un 3,5 %. Según otro estudio [98] en el que se analiza la tasa de respuesta de encuestas utilizando emails, el uso de recordatorios es uno de los principales factores para aumentar dicha tasa. En general, el uso de emails como recordatorio suele aumentar la cantidad de usuarios que responden las encuestas en un 25 %. Además, el uso de varios recordatorios tiene una mayor influencia positiva en la tasa de respuesta que el uso de un único recordatorio. Sin embargo, el uso de un excesivo número de recordatorios tiende a un estancamiento en el número de respuestas, ya que se presenta una minusvaloración de la encuesta.

Esta opción, que no se había planteado inicialmente, fue tomada en cuenta debido a la baja tasa de respuestas obtenida a falta de un mes

TASA DE MODIFICACIÓN DE LA POBLACIÓN MUESTRAL: Aumento o disminución de la población final con respecto a la población original, por la inclusión de nuevos individuos.

para el cierre de las encuestas. Para intentar aumentar dicha tasa, se enviaron 2 recordatorios, el primero remitido a falta de un mes para el cierre de la encuesta y el segundo recordatorio se envió a falta de 15 días. En la primera de esas fechas, utilizando la base de datos del sistema Web de encuestas en la que se tenía constancia de las cartas de presentación enviadas previamente, se remitieron 5 tipos de recordatorios, en función de la respuesta de cada encuestado con respecto a la encuesta:

- Si habían contestado a los dos cuestionarios era una carta de agradecimiento. Mediante la cual se ha pretendido mantener activo el interés por el estudio, para reclutar más encuestados.
- Si sólo habían rellenado uno de los cuestionarios, se les agradecía el hecho de haber rellenado uno y se les recordaba que la encuesta constaba de dos cuestionarios.
- Si no habían llegado a contestar ningún cuestionario, se les agradecía el hecho de haberse registrado y se les informaba que había 2 cuestionarios por contestar.
- Si no habían contestado ninguna pregunta, se les agradecía el hecho de haberse registrado y se les preguntaba si existía algún problema técnico que le impidiese rellenar los cuestionarios para intentar solventarlo lo antes posible.
- Si no se habían registrado, se les solicitaba que se registrasen y rellenasen los cuestionarios.

En la segunda remesa de recordatorios se enviaron 7 tipos de recordatorios diferentes, en función de la respuesta de cada usuario con respecto a los cuestionarios y si se le había mandado recordatorio de ese mismo tipo en el envío anterior. Para apremiar la respuesta, en este envío, se repitió varias veces y se puso en mayúsculas tanto el tiempo restante ("15 DÍAS") como la fecha exacta de cierre de la encuesta. Los diferentes tipos de envío se describen a continuación:

- Dos cuestionarios contestados, sin envío de este tipo en la remesa anterior: es decir, habían rellenado la encuesta en el periodo transcurrido entre el primer recordatorio y el segundo recordatorio. Se les enviaba un email de agradecimiento idéntico al que se les había mandado a los que habían contestado los dos cuestionarios en el envío anterior.
- Un cuestionario contestado sin recordatorio de este tipo en la remesa anterior: por tanto, habían rellenado una encuesta en el periodo entre los dos reenvíos de recordatorios. Se les agradecía el hecho de haber rellenado un cuestionario y se les recordaba que la encuesta constaba de dos.
- Un cuestionario contestado con recordatorio previo del mismo tipo en el envío anterior: con lo que no habían continuado rellenando o no habían llegado a finalizar el segundo cuestionario. Se les indicaba que era la segunda vez que se le había enviado un recordatorio al encuestado y se le urgía a que finalizase la encuesta.

- Ningún cuestionario contestado sin envío de este tipo de recordatorio en el envío anterior: estos casos son de usuarios que en el envío anterior no estaban registrados o no habían contestado ninguna pregunta y que en este envío tenían alguna pregunta contestada. Se les agradecía el hecho de haberse registrado y se les informaba que había 2 encuestas por contestar.
- Ningún cuestionario contestado con un recordatorio previo de este tipo en el envío anterior: estos casos son de usuarios que continúan en este envío con alguna pregunta contestada pero sin llegar a completar al menos una encuesta. Se les indicaba que se les había enviado un email previamente y se les apremiaba a finalizar los dos cuestionarios.
- Si no habían contestado ninguna pregunta, se les agradecía el hecho de haberse registrado y se les preguntaba si existía algún problema técnico que le impidiese rellenar los cuestionarios para intentar solventarlo lo antes posible. En este caso no se tuvo en cuenta si se había enviado algún email previo.
- Si no se habían registrado, se les solicitaba que se registrasen y rellenasen los cuestionarios. En este caso, tampoco se tuvo en cuenta si ya se había enviado un recordatorio anterior.

En la Tab. 10.1 se muestra una tabla con la evolución del número de usuarios registrados en el sistema Web de encuestas, el número de encuestados que habían contestado, respectivamente, las 2 encuestas, sólo 1 de las dos encuestas, algunas preguntas de una de las encuestas o ninguna pregunta (sólo se habían registrado en el sistema pero no habían contestado ninguna pregunta, marcado en la tabla en la columna S/R). Se puede observar cómo el aumento de la respuesta ha sido muy importante tras el envío de los recordatorios. En particular, es relevante el impacto que sufre el número de usuarios registrados, pasando de 48 a 56 en sólo 3 días y a 65 en 14 días, y el número de usuarios que finalizaron los 2 cuestionarios, que aumentó de 12 a 15 en los 3 primeros días y a 27 en 2 semanas. Hay que entender que el aumento de encuestas finalizadas completamente es más suave a lo largo de un tiempo mayor que el incremento de nuevos usuarios registrados, ya que la contestación de los cuestionarios es un proceso más largo temporalmente y por tanto, más gradual. En cualquier caso, se puede observar que el envío de estos recordatorios supuso un cambio de tendencia positiva muy acusada.

10.4.3 Tasa de respuesta

Como se ha indicado con anterioridad, al utilizar la técnica de muestreo de *bola de nieve*, no es posible obtener una tasa de respuesta exacta, sino sólo una estimación de la misma a partir del número de cartas de presentación enviadas originalmente. Algunos usuarios al remitir la carta de presentación a sus listas de contactos, han puesto en copia al encuestador, por lo que en dichos casos, se tiene el número adicional de individuos a los que se le ha enviado la carta de presentación, que incrementa la población muestral original. En los casos en los que

FECHA	USUARIOS	CUESTIONARIOS CONTESTADOS			
	REGISTRADOS	S/R	0	1	2
19-02-2010	39	6	17	9	7
05-03-2010	48	4	23	9	12
08-03-2010	56	4	25	12	15
19-03-2010	65	1	22	15	27
22-03-2010	68	1	22	16	29
06-04-2010	70	2	18	13	37

Tabla 10.1
Agrupado por fechas, usuarios registrados, usuarios que han contestado los 2 cuestionarios, sólo 1 de ellos, sólo algunas preguntas pero ningún cuestionario completo, ninguna pregunta (S/R = Sin respuestas).

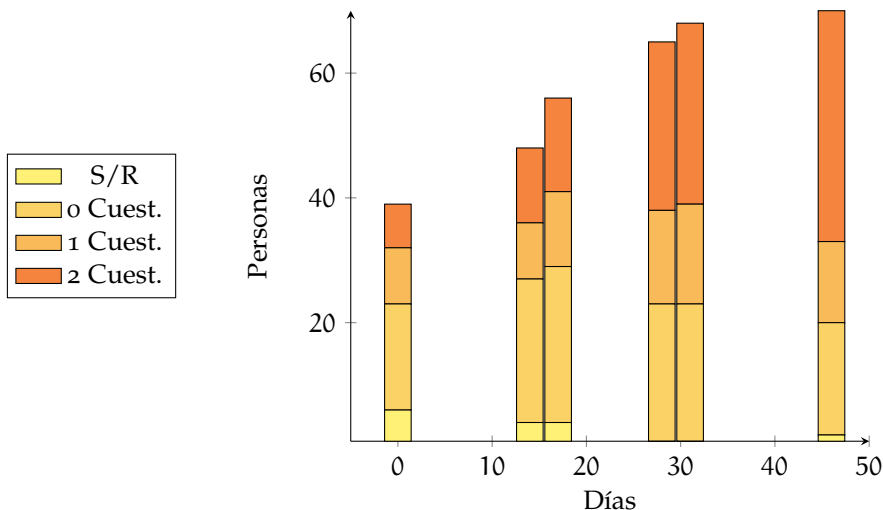


Figura 10.1
Gráfica de respuesta de encuestas

los usuarios pertenecen a algún grupo común, las listas de contactos de los usuarios tienen un cierto grado de solapamiento, normalmente no completo. En dichos casos, no se presenta un aumento de la muestra poblacional, aunque de manera indirecta se consigue un cierto efecto recordatorio, sobre todo si el reenvío de la carta de presentación al grupo común por parte de cada uno de los usuarios se hace con un intervalo de tiempo amplio. Estos recordatorios, aunque deseables, son indirectos e incontrolables, ya que no es el propio encuestador el que realiza su envío, y por tanto, no puede determinar ni el texto de los recordatorios, ni las fechas en las que se envían, ni sobre qué individuos se realiza.

Partiendo del envío de la carta de presentación a 26 individuos, se tiene constancia del reenvío de dicha carta de presentación por 7 de ellos, por lo que la **TASA DE DIFUSIÓN INICIAL** ha sido de 26.9. Estos 7 reenvíos hicieron llegar la carta de presentación aproximadamente a 160–175 individuos más, por lo que se tuvo una muestra poblacional total de 185–195 individuos. Por tanto, la **TASA DE AMPLIACIÓN DE LA POBLACIÓN** muestral estuvo entorno al 269.2%, lo que es un grado de aumento de la población muestral bastante importante. Estos datos indican que no toda la población original fue escogida correctamente, ya que hubo pocos individuos de dicha población original que actuaron como vectores de propagación de la encuesta. Sin embargo, aquellos

TASA DE DIFUSIÓN INICIAL: Medida de la cantidad de población inicial que ha colaborado en la ampliación de la población final.

Tabla 10.2
Tasa de Respuesta respecto a la muestra poblacional estimada (GENERAL) y respecto al número de usuarios registrados (ADAPTADA) según usuarios (Registrados, Encuestas Completas o Algún Cuestionario).

	NÚMERO DE RESPUESTAS	TASA DE RESPUESTA	
		GENERAL	ADAPTADA
Usuarios Registrados	70	36.8 %	100.0 %
Encuestas completas	37	19.5 %	52.9 %
Algún cuestionario	50	26.3 %	71.4 %

que reenviaron la encuesta fueron muy productivos.

Tomando una estimación del tamaño de la muestra poblacional de 190 individuos, a continuación se van a proporcionar varias medidas de la tasa de respuesta en función de los usuarios registrados, del número de cuestionarios contestados y del número de individuos que hayan contestado completamente los dos cuestionarios que la componen. En la Tab. 10.2 se muestran las diversas medidas de tasa de respuesta con respecto a la muestra poblacional completa estimada, relacionándola con los usuarios registrados en el sistema Web de encuestas, los usuarios que han contestado completamente la encuesta (las 2 encuestas) y los usuarios que han contestado al menos 1 de los dos cuestionarios que componen la encuesta. También se ha incluido una tasa de respuesta, denominada TASA DE RESPUESTA ADAPTADA, en la que las respuestas se comparan con el número de usuarios registrados.

10.5 RESULTADOS DEL PROCESO DE EVALUACIÓN

Como se ha comentado con anterioridad, el proceso de diseño y de respuesta de los cuestionarios presenta suficiente interés desde el punto de vista científico como para incluir tanto el análisis de los cuestionarios tras su respuesta por parte de los encuestados, como la recopilación de los comentarios a las encuestas.

10.5.1 Respuestas cualitativas

Cada vez que un usuario registrado en el sistema Web finalizaba los dos cuestionarios de la encuesta y, en general, al finalizar el periodo de respuestas de los cuestionarios, cada usuario recibía un email de agradecimiento, en el que además se le solicitaba diera su opinión sobre el propio proceso. Estas respuestas contrastan con las respuestas a las preguntas planteadas en la encuesta, ya que estas últimas pretenden proporcionar un resultado numérico o cuantitativo. Por el contrario, las primeras proporcionan indicios de tipo cualitativo, que son difícilmente comparables, pero mucho más ricas desde el punto de vista subjetivo. Aproximadamente un 10 % de los encuestados que finalizaron completamente los dos cuestionarios (7 usuarios) contestaron, coincidiendo en la mayoría de opiniones, que se pueden resumir en los siguientes puntos:

- La repetición de preguntas de mapeado espacial y temporal, en la que se utiliza siempre la misma imagen base (Fig. 7.17) y en

la que siempre se pregunta por el mismo elemento (*el coche en sombras*), hace que una vez que se identifique la posición de dicho elemento, se pueda identificar en todas las demás repeticiones de dicha pregunta, aún cuando se cambie el algoritmo de extracción de contornos.

- El uso de preguntas de puntuación sin repetición plantea un cierto grado de incomodidad, ya que no permite tanta flexibilidad y obliga a puntuar con mucho mayor grado de precisión entre las diferentes imágenes de respuesta.
- A pesar de la explicación de la finalidad de la encuesta, los usuarios tienen dudas acerca de para qué pueden servir sus respuestas, por lo que, en general, les falta incentivo para continuar con los cuestionarios.
- El uso de dos cuestionarios con las mismas preguntas desconcierta.
- Los cuestionarios son excesivamente exigentes y largos de responder.

Algunos de los usuarios, además de indicar los problemas o inconvenientes encontrados, también proporcionaron ideas para solventarlos o limitarlos en cierta medida. Por ejemplo, para las preguntas de mapeado espacial y temporal, se planteó la utilización de imágenes base similares, que presentaran los mismos objetos, pero modificando la posición de los mismos en la escena, alterando también las sombras proyectadas y buscando en cada uno de ellos un objeto diferente. Esta propuesta se observó interesante, pero abre un estudio colateral para determinar cómo influyen todas esas variables en las elecciones de los usuarios, por lo que su adopción como tipo de pregunta debería estar supeditada a que dichos estudios mostraran un alto grado de independencia en las respuestas.

10.5.2 Metodología de las respuestas

Algunas respuestas cualitativas de los encuestados también incluyeron una breve descripción de cómo habían procedido al contestar las preguntas de los cuestionarios. Estas respuestas, unidas a observaciones presenciales no invasivas de la forma en la que los usuarios han contestado las encuestas, han permitido extraer algunas pautas metodológicas en el proceso de respuesta. No se puede afirmar que exista una metodología única, sino que cada usuario tiene su manera única y personal de responder las encuestas, aunque dentro de esas maneras propias hay algunos pasos que son compartidos por a muchos encuestados.

Existen algunos comportamientos comunes en muchos de los encuestados, entre los que cabe destacar:

- En las preguntas de selección, muchas imágenes se descartan sin necesidad de abrir la imagen, es decir, se evalúan sólo utilizando las imágenes en miniatura. Sólo aquellas que presentan dudas se abren para verlas en tamaño maximizado.

- Muchos usuarios, antes de puntuar las imágenes, maximizan cada una de ellas y las observan detenidamente, comparándolas, aunque sin puntuarlas en ese momento. La puntuación efectiva se realiza en una segunda o tercera pasada.
- A la hora de puntuar las imágenes de contornos, hay dos tipologías de encuestados:
 - A. Usuarios que contestan en orden de aparición de las imágenes de respuesta de izquierda a derecha y de arriba a abajo.
 - B. Usuarios que ordenan las imágenes de respuesta en función de la calidad, bien de peor a mejor o al revés, y puntúan estas imágenes ordenadamente, asignando puntuaciones muy bajas (o muy altas) a las primeras, que actúan de imágenes de referencia para puntuar el resto.

En general, aunque existen algunos comportamientos comunes en muchos usuarios, se puede afirmar que la metodología que cada encuestado desarrolla para contestar los cuestionarios, es única, diferente y adaptada por cada individuo para facilitarle la comparación entre imágenes y el proceso de contestación de las preguntas.

10.6 ANÁLISIS PRELIMINAR DE LOS CUESTIONARIOS

Antes de analizar de manera global las respuestas proporcionadas por los encuestados, es conveniente realizar algunos estudios preliminares que permitan determinar la calidad de los cuestionarios desarrollados para asegurar que las inferencias y deducciones que se realicen a partir de los datos obtenidos de las encuestas sean estadísticamente relevantes. Las recomendaciones de diseño expuestas en el Cap. 8 no garantizan que los cuestionarios sean perfectos, por lo que, hay que realizar un estudio de qué y cómo han respondido los encuestados.

10.6.1 *Análisis de la población*

Se considera como población de la encuesta sólo aquellos encuestados que hayan contestado todas las preguntas de los dos cuestionarios, es decir, se parte de una población de 37 encuestados. En el paso de registro de los usuarios en el sistema Web de encuestas, cada encuestado ha tenido que rellenar un perfil que ha permitido catalogar su grado de visión, puntuando de 0 (no le afecta) a 10 (le afecta completamente) una serie de problemas visuales:

- Miopía
- Hipermetropía
- Astigmatismo
- Retinosis pigmentaria
- Miopía magna
- Degeneración macular

- Cataratas
- Desprendimiento de retina
- Pérdida de visión lateral

Analizando el perfil de visión de cada encuestado se puede hacer una distinción de dos grupos: Encuestados con un grado de VISIÓN ESTÁNDAR y encuestados con BAJA VISIÓN. Se ha considerado que un encuestado sufre de BAJA VISIÓN si presenta alguna de las siguientes características:

- Puntuación superior a 7 en 3 o más problemas visuales.
- Puntuación superior a 5 en al menos uno, o superior a 3 en dos o más, de los siguientes problemas visuales:
 - Retinosis pigmentaria
 - Miopía magna
 - Degeneración macular
 - Cataratas
 - Desprendimiento de retina
 - Pérdida de visión lateral

Utilizando esta catalogación se ha podido establecer que en la población de la encuesta hay 5 encuestados con problemas de BAJA VISIÓN y 32 usuarios que tienen un nivel de VISIÓN ESTÁNDAR.

10.6.2 Muestreo a posteriori

Como paso previo también es necesario analizar la viabilidad de todos los individuos que han respondido a los cuestionarios, es el muestreo *a posteriori* indicado con anterioridad en la Sec. 10.4. En esta fase se detectan **OUTLIERS** como respuestas de encuestados que se alejan en exceso de la respuesta media. Ésta se calcula como la media de las respuestas de cada usuario para cada una de las preguntas de la encuesta. Utilizando esta respuesta media, se calcula la diferencia de las respuestas de cada encuestado, que puede ser considerado como medida del error de cada encuestado para cada pregunta. Según el **TEOREMA DEL LÍMITE CENTRAL** al tomar 203 muestras (una por cada imagen de respuesta de cada pregunta de los dos cuestionarios) de los 37 encuestados, la respuesta media sigue una distribución normal con media $\mu = -7.004 \cdot 10^{-18}$ y varianza $\sigma^2 = 22.78$. Mediante el test de Normalidad de *Shapiro-Wilk* [97] se comprueba que la respuesta media sigue una distribución normal con un nivel de significación del 95 %, ya que se obtiene un valor de probabilidad $p = 0.7394 > 0.05$.

Teniendo la garantía de que la respuesta media de los encuestados sigue una distribución normal, se puede aplicar el método de *Grubb* [45] para la detección de **OUTLIERS**, cuyo resultado se muestra en la Tab. 10.3. Para un nivel de significación del 95 %, este método proporciona el valor crítico del estadístico $Z = 3.0026$, que permite estimar para cada encuestado la existencia de valores extremos. En este caso, el método de *Grubb* no detecta ningún **OUTLIER** entre las respuestas de los

OUTLIERS:
Respuestas extremas que no están dentro de un comportamiento normal y que deberían ser descartadas.

TEOREMA DEL LÍMITE CENTRAL: Si se toman un número elevado (n) de muestras aleatorias independientes de una población con media μ y varianza σ^2 , la distribución muestral de la suma de dichas muestras sigue una distribución normal de media μ y varianza $\frac{\sigma^2}{n}$.

encuestados. Por lo tanto, se pueden aceptar a todos los encuestados dentro de la población para el análisis definitivo de las respuestas de los cuestionarios.

10.6.3 *Análisis de frecuencia de cada pregunta*

Un indicador de la calidad de una encuesta proviene del análisis de frecuencia de cada pregunta, en el que se calcula cuántos encuestados han contestado a cada valoración para cada pregunta. Gracias a este análisis se puede encontrar qué preguntas son votadas de manera uniforme por todos los encuestados y descartarlas, en su caso, ya que dichas preguntas no discriminan ni proporcionan mayor información, puesto que todos los encuestados han puntuado esas preguntas de igual manera. En los dos cuestionarios que forman la presente encuesta, cada pregunta está compuesta por una o varias imágenes de respuesta. En esta sección, se analizan de manera individual cada una de estas imágenes de respuesta y luego, se estudian las preguntas como conglomerado de dichas imágenes de respuesta.

Se han encontrado 2 imágenes de respuesta para las que todos los encuestados han puntuado con la máxima puntuación (10) y 14 imágenes de respuesta en las que todos los encuestados han puntuado con la puntuación mínima (0). Sin embargo, sólo una pregunta al completo ha sido puntuada de manera uniforme para todas las imágenes de respuesta que la componen: todos los encuestados puntuaron con un 0 a todas las imágenes de respuesta que componen la pregunta número 7 del primer cuestionario. Esto quiere decir, que se podría eliminar dicha pregunta del cuestionario sin producir un efecto en la capacidad de discriminación de la encuesta. Teniendo en cuenta que los dos cuestionarios presentan un total de 203 imágenes de respuesta agrupadas en 44 preguntas diferentes, el efecto del número de imágenes de respuesta con variación de frecuencia nula es únicamente de 7.88% y el efecto en el número de preguntas con variación de frecuencia nula es de un 2.27%. Estos valores tan bajos, permiten afirmar que la gran mayoría de preguntas e imágenes de respuesta incluidos en los cuestionarios son discriminantes y su inclusión en la encuesta es pertinente.

ENCUESTADO	PUNTUACIÓN MEDIA	ESTADÍSTICO Z	¿OUTLIER?
1	0.70	0.895	No
2	-0.42	0.536	No
3	-0.67	0.856	No
4	-0.37	0.473	No
5	-0.78	0.997	No
6	-1.09	1.393	No
7	0.29	0.371	No
8	-0.24	0.306	No
9	-1.58	2.019	No
10	-0.01	0.012	No
11	0.55	0.703	No
12	-0.60	0.767	No
13	0.46	0.588	No
14	-0.06	0.076	No
15	-0.05	0.064	No
16	0.13	0.167	No
17	-0.71	0.907	No
18	-1.31	1.674	No
19	0.60	0.767	No
20	-1.10	1.406	No
21	0.80	1.023	No
22	1.34	1.713	No
23	0.82	1.048	No
24	-0.41	0.524	No
25	0.92	1.176	No
26	-0.54	0.690	No
27	-0.40	0.511	No
28	1.17	1.496	No
29	-0.32	0.409	No
30	1.08	1.381	No
31	0.06	0.077	No
32	0.34	0.435	No
33	0.36	0.460	No
34	-0.06	0.076	No
35	1.41	1.802	No
36	-1.20	1.533	No
37	0.88	1.125	No

Tabla 10.3
Detección de outliers
mediante el método
de *Grubb*.

10.7 ANÁLISIS DE LOS RESULTADOS

Esta encuesta pretende obtener la valoración subjetiva de seis métodos trabajando con dos conjuntos de parámetros, bajo las restricciones indicadas en el Cap. 7. En la Tab. 10.4 se muestran las puntuaciones obtenidas en la encuesta sin aplicar ningún tipo de ponderación por pregunta. En la misma se puede observar que todos los métodos LIP obtienen puntuaciones mejores que sus correspondientes métodos No LIP. En particular, ningún método No LIP consigue aprobar, es decir, obtener una puntuación superior al 5, salvo el método CANNY-80 (que representa al método *Canny* con un umbral máximo del 80% de los píxeles de la imagen marcados como contornos) que obtiene un 5.68. Sin embargo, los métodos LIP obtienen puntuaciones subjetivas superiores al 5 (salvo, los basados en el *Laplaciano de Gaussianas*, LIPLoG-95 y LIPLoG-120, que obtienen valores cercanos al 5, sin superarlo). En la Fig. 10.2 se comparan de manera gráfica los resultados subjetivos obtenidos. En dicha imagen, se han representado con barras las puntuaciones subjetivas obtenidas para cada método, donde la barra más a la izquierda de cada bloque representa el método en su versión No LIP y la barra a su derecha el mismo método en su versión LIP.

10.7.1 Análisis por grado de visión de los encuestados

Los resultados obtenidos han sido proporcionados por una población que no es homogénea, ya que la población que ha respondido esta encuesta está integrada tanto por encuestados con BAJA VISIÓN como por personas con un grado de VISIÓN ESTÁNDAR. En particular, han respondido los dos cuestionarios 5 personas que han sido categorizados como individuos con BAJA VISIÓN, mientras que los 32 usuarios restantes poseen un grado de VISIÓN ESTÁNDAR. En esta sección se muestran los resultados según el grado de visión de los encuestados, agrupados en los dos tipos anteriormente citados.

En la Tab. 10.5 se muestran las puntuaciones subjetivas obtenidas para cada método bajo tres referencias: según la población completa (primera columna), teniendo en cuenta sólo las respuestas de los usuarios con BAJA VISIÓN (segunda columna) y tomando únicamente a los encuestados con un grado de VISIÓN ESTÁNDAR (última columna). Como se puede observar, aunque las puntuaciones no tienen el mismo valor, sí que muestran la misma tendencia; es decir, los métodos peor valorados son igualmente peor valorados por todos los tipos de usuarios y los métodos mejor valorados obtienen mejores valoraciones en todos los casos, independientemente del grado de visión de los encuestados. Para facilitar la comparación se incluye la Fig. 10.3, en la que se representa mediante barras la puntuación media de cada método. De igual manera que en la Fig. 10.2, los métodos se han agrupado por parejas, en las que la barra izquierda de cada par representa la versión No LIP del método en cuestión y la barra derecha, la versión LIP de dicho método. A su vez, en la Fig. 10.3, cada barra está dividida en 3 cortes. La marca inferior indica la media de los encuestados con BAJA VISIÓN, el corte intermedio representa la media teniendo en cuenta la población

MÉTODO	PUNTUACIÓN MEDIA	DESVIACIÓN TÍPICA
SOBEL-80	2.04	0.72
SOBEL-90	2.31	0.58
CANNY-80	5.68	0.93
CANNY-90	3.70	0.93
LoG-95	3.33	0.93
LoG-120	2.77	0.67
LIPSOBEL-80	5.63	1.01
LIPSOBEL-90	6.33	1.52
LIPCANNY-80	6.87	1.09
LIPCANNY-90	7.76	0.91
LIPLoG-95	4.53	1.23
LIPLoG-120	4.17	1.33

Tabla 10.4
Puntuación subjetiva
de cada método.

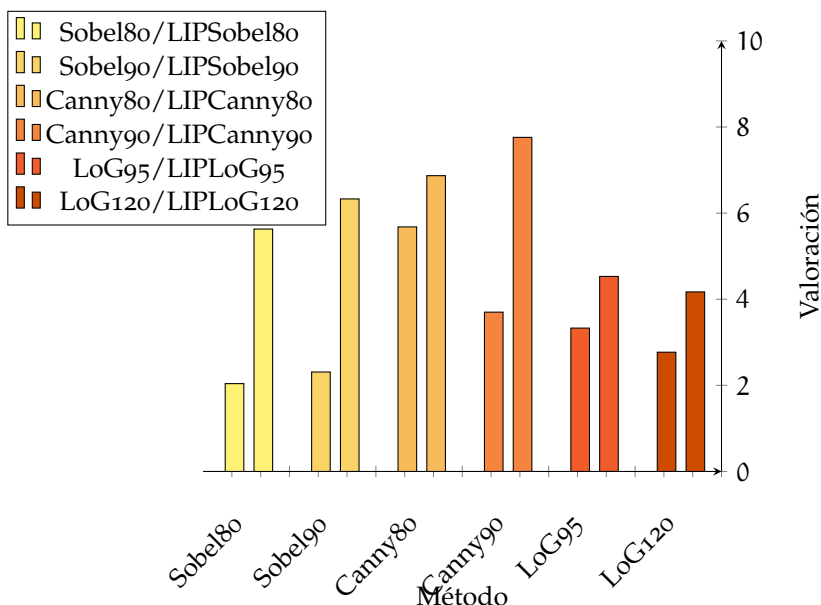


Figura 10.2
Valoración subjetiva
de los métodos

COMPLETA y el tope superior de cada barra coincide con la respuesta de los usuarios con VISIÓN ESTÁNDAR.

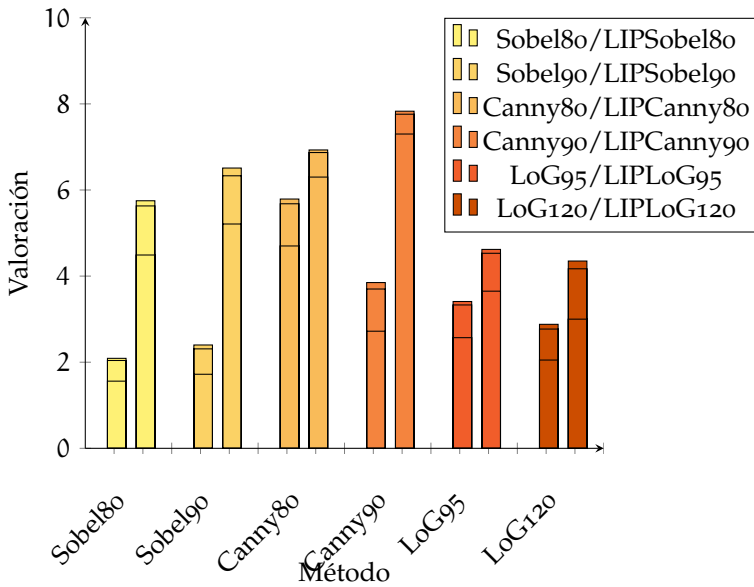
Correlación de las respuestas de los tipos de encuestados

Las respuestas que proporcionan los usuarios, categorizados según su grado de visión, muestran un elevado nivel de coherencia en los diferentes métodos. Para comprobar esta coherencia, se ha calculado la tabla de correlaciones cruzadas entre las respuestas de los diferentes tipos de usuarios, mostrada en la Tab. 10.6. La correlación calculada se ha obtenido utilizando el coeficiente *r* de Pearson, por lo que los valores pueden oscilar entre 0 (no existe relación alguna entre los dos conjuntos de variables) y 1 (un conjunto es linealmente dependiente del otro). La relación más interesante de dicha tabla es la que presentan las res-

Tabla 10.5
Puntuación subjetiva de cada método según el grado de visión.

MÉTODO	POBLACIÓN COMPLETA	BAJA VISIÓN	VISIÓN ESTÁNDAR
SOBEL-80	2.04	1.56	2.09
SOBEL-90	2.31	1.72	2.40
CANNY-80	5.68	4.70	5.79
CANNY-90	3.70	2.72	3.85
LoG-95	3.33	2.57	3.41
LoG-120	2.77	2.05	2.88
LIPSOBEL-80	5.63	4.49	5.75
LIPSOBEL-90	6.33	5.21	6.51
LIPCANNY-80	6.87	6.30	6.93
LIPCANNY-90	7.76	7.30	7.83
LIPLoG-95	4.53	3.65	4.62
LIPLoG-120	4.17	3.00	4.35

Figura 10.3
Valoración subjetiva de los métodos



puestas de los usuarios con BAJA VISIÓN y los encuestados con VISIÓN ESTÁNDAR, que muestra un grado de correlación de 0.9874. Observando dicha tabla, se puede comprobar que este valor es el más bajo de toda la tabla de correlaciones cruzadas, por lo que la correlación entre dos grupos cualesquiera siempre se encuentra por encima de 0.98. Este valor mínimo tan elevado, permite afirmar que todas las respuestas están altamente relacionadas entre sí.

Proporción de las respuestas de los tipos de encuestados

Gracias al coeficiente r de Pearson calculado, se ha podido determinar estadísticamente que las respuestas de los diferentes tipos de usua-

COEFICIENTE DE CORRELACIÓN	POBLACIÓN COMPLETA	BAJA VISIÓN	VISIÓN ESTÁNDAR
POBLACIÓN COMPLETA	1	0.9903	0.9998
BAJA VISIÓN	0.9903	1	0.9874
VISIÓN ESTÁNDAR	0.9998	0.9874	1

Tabla 10.6
Grado de correlación (coeficiente r de Pearson) entre las respuestas de los distintos tipos de usuarios según su grado de visión.

rios están relacionadas, lo cual no significa que las respuestas de los distintos tipos de usuarios sean exactamente iguales. De hecho, al estudiar la Fig. 10.3 se observa que, para todos los métodos analizados, las respuestas de los usuarios con BAJA VISIÓN son algo inferiores en puntuación a las respuestas de los usuarios con VISIÓN ESTÁNDAR. Para estimar si la bajada en las puntuaciones de los encuestados con BAJA VISIÓN sigue una proporción común en todos los métodos o si cada método tiene una proporción diferente, se debe realizar un análisis más pormenorizado.

En la Tab. 10.7 se muestra la proporción del valor medio de puntuación para cada método evaluado según el tipo de usuario. Los resultados mostrados se han obtenido aplicando (10.1).

$$p = \frac{V_A}{V_B} \quad (10.1)$$

Donde V_A representa el valor medio del primer conjunto de encuestados comparados y V_B representa el valor medio del segundo grupo en la comparación.

La columna BV-C compara los valores medios de los pacientes con BAJA VISIÓN con respecto a los valores medios tomando la población COMPLETA. La segunda columna, VE-C, compara los valores medios de las respuestas de los encuestados con grado de VISIÓN ESTÁNDAR con respecto a las puntuaciones medias tomando la población COMPLETA. La última columna, VE-BV, compara las respuestas de los encuestados con grado de VISIÓN ESTÁNDAR con los resultados medios de los usuarios con BAJA VISIÓN. Dicha columna es la que proporciona una información más relevante, ya que las dos poblaciones comparadas son dos conjuntos que no presentan solapamientos. Se puede observar que las respuestas de los encuestados con un grado de VISIÓN ESTÁNDAR son entre un 10 % y un 45 % superiores a las respuestas de los encuestados que sufren BAJA VISIÓN.

Se ha incluido la Tab. 10.8 en la que se muestran los valores medios y las desviaciones típicas de las proporciones calculadas en la Tab. 10.7, agrupando los métodos en los dos grandes bloques analizados, métodos No LIP y métodos LIP. El primer grupo incluye los métodos SOBEL-80, SOBEL-90, CANNY-80, CANNY-90, LOG-95 y LOG-120; mientras que el segundo grupo está compuesto por los métodos LIPSOBEL-80, LIPSOBEL-90, LIPCANNY-80, LIPCANNY-90, LIPLoG-95 y LIPLoG-120. También se ha incluido el valor medio y la desviación típica considerando todos los métodos.

Tabla 10.7
Relación proporcional entre las puntuaciones subjetivas medias de cada método según los distintos grados de visión. C: Población completa. BV: Individuos con Baja Visión. VE: Individuos con visión estándar.

MÉTODO	RELACIÓN PROPORCIONAL		
	BV-C	VE-C	VE-BV
SOBEL-80	0.77	1.03	1.34
SOBEL-90	0.75	1.04	1.40
CANNY-80	0.83	1.02	1.23
CANNY-90	0.74	1.04	1.42
LOG-95	0.77	1.03	1.33
LOG-120	0.74	1.04	1.40
LIPSOBEL-80	0.80	1.02	1.28
LIPSOBEL-90	0.82	1.03	1.25
LIPCANNY-80	0.92	1.01	1.10
LIPCANNY-90	0.94	1.01	1.07
LIPLoG-95	0.81	1.02	1.27
LIPLoG-120	0.72	1.04	1.45

Tabla 10.8
Media y Desviación Típica de los métodos LIP y No LIP entre las respuestas de los distintos tipos de usuarios según su grado de visión.

MÉTODO	RELACIÓN PROPORCIONAL		
	BV-C	VE-C	VE-BV
MEDIA NO LIP	0.7645	1.0316	1.3521
DESV. TÍP. NO LIP	0.0343	0.0814	0.0694
MEDIA LIP	0.8342	1.0221	1.2362
DESV. TÍP. LIP	0.0814	0.0129	0.0136
MEDIA TOTAL	0.7994	1.0269	1.2941
DESV. TÍP. TOTAL	0.0698	0.0120	0.1195

En la Tab. 10.8 se puede observar que, para los métodos NO LIP, las respuestas de los encuestados con BAJA VISIÓN muestran una proporción más baja (son inferiores en algo más de un 35 %) frente a las respuestas de los encuestados con un grado de VISIÓN ESTÁNDAR, mostrando además una mayor desviación típica, que se ha de interpretar como una mayor variabilidad en las respuestas. Por el contrario, las respuestas de los encuestados ante los métodos LIP, muestran un comportamiento más estable (ya que presentan una desviación típica en torno al 1 %) y con valores medios de respuesta más cercanos entre los encuestados con BAJA VISIÓN y los usuarios con VISIÓN ESTÁNDAR, puesto que las respuestas de los primeros son solamente un 23 % inferiores a los valores medios de las respuestas de los segundos.

Regresión lineal de las respuestas de los tipos de encuestados

Para comprobar que la relación de las respuestas entre los dos grupos de encuestados es lineal, se ha calculado la regresión lineal de las respuestas de los encuestados con BAJA VISIÓN para cada uno de los métodos evaluados con respecto a las respuestas de los encuestados con VISIÓN ESTÁNDAR para los mismos métodos. El ajuste de los

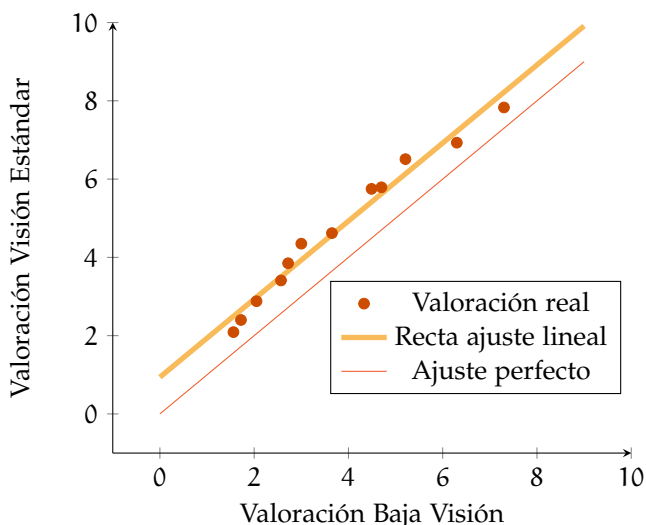


Figura 10.4
Gráfica de ajuste
lineal de las respues-
tas de los encuesta-
dos según grado de
visión.

puntos a la recta se ha realizado mediante ajuste lineal de mínimos cuadrados utilizando la técnica de *Descomposición QR*. Este cálculo ha permitido obtener el ajuste mostrado en (10.2), con un error cuadrático acumulado total de 0.9611 (con error cuadrático medio de 0.0801).

$$Y_{VE} = 0.9970 \cdot X_{BV} + 0.9397 \quad (10.2)$$

Donde Y_{VE} es el ajuste lineal correspondiente con las respuestas de los encuestados con VISIÓN ESTÁNDAR y X_{BV} representa el valor de las respuestas de los encuestados con BAJA VISIÓN.

En la Fig. 10.4 se han representado los pares de puntuaciones para cada método donde el eje X está asociado a las respuestas de los encuestados con BAJA VISIÓN, mientras que el eje Y está asociado a los valores de los encuestados con VISIÓN ESTÁNDAR. También se ha representado la recta de regresión obtenida a partir de (10.2), que ajusta linealmente el conjunto de puntos. Para facilitar la comparación, se ha incluido la recta de ajuste perfecto, $y = x$.

Gracias a los valores de ajuste obtenidos se puede afirmar que, en un hipotético caso medio, los valores que proporcionan las respuestas de los encuestados con BAJA VISIÓN son prácticamente iguales (0.997) a los valores de las respuestas de los encuestados con VISIÓN ESTÁNDAR, salvo que las respuestas de estos últimos están sobreestimadas en aproximadamente 1 punto (0.9397).

10.8 VALIDACIÓN DE LA ENCUESTA

Una encuesta se considera una herramienta que mide una determinada variable. Esta herramienta es útil como medida, si realmente es capaz de medir de manera fiable la variable para la que se ha diseñado. Se entiende que una encuesta es fiable si proporciona las mismas mediciones cada vez que se aplica a los mismos encuestados en similares circunstancias, o si proporciona mediciones análogas frente a distintos usuarios con similares características personales.

Mediante la validación de los cuestionarios, los investigadores garantizan que las encuestas miden realmente las variables que se pretendían medir cuando se diseñaron dichos cuestionarios. La validez no garantiza que una encuesta sea repetible, ni que se pueda extrapolar a otros campos o poblaciones y en algunos casos. Tampoco es capaz de establecer el grado de fiabilidad de las medidas. Tradicionalmente se han utilizado tres tipos de validaciones:

1. Validez de criterio.
2. Validez de constructo.
3. Validez de contenido.

A continuación, se muestran los resultados del estudio de validación de los cuestionarios que se presentan en esta Tesis Doctoral. Este análisis garantiza que la encuesta que se ha diseñado mide realmente la calidad percibida por los usuarios de diferentes algoritmos de extracción de contornos.

10.8.1 *Validez de criterio*

Cuando existe un criterio de medición preestablecido o cuando existe un patrón GOLD STANDARD que permita conocer con exactitud el valor exacto de la variable que se pretende medir mediante el cuestionario, se puede determinar el grado de correlación entre las dos medidas. En caso de coincidencia, se garantiza la validez de criterio. Para estos cuestionarios, no es posible realizar una validez de criterio con un cierto grado de fiabilidad, ya que aunque existen otros mecanismos de tipo objetivo que permiten proporcionar una medida numérica de la calidad de los contornos, estos métodos no están testeados con respecto al grado de calidad percibida por parte de los usuarios.

10.8.2 *Validez de constructo*

La validez de constructo pretende comprobar si las respuestas reales de los encuestados se agrupan en los mismos dominios de las variables que se plantearon a nivel teórico al diseñar los cuestionarios. En esta investigación, no se ha podido obtener ningún indicio concluyente que permita demostrar la validez de constructo para los cuestionarios diseñados. De manera habitual, la validez de constructo se suele obtener mediante el método PCA, sin embargo, no ha sido posible aplicarlo ya que no se cumplen varios de los requisitos que se imponen a los datos experimentales para poder aplicar PCA:

- La matriz de correlaciones debe tener varias correlaciones con valor superior a 0.5.
- El número de encuestados debe ser, al menos, de 50 (preferible, 100).
- El ratio de encuestados a variables debe ser de 5 a 1.

MÉTODO	VALORACIÓN
EXPERTO1	8
EXPERTO2	7
EXPERTO3	9
EXPERTO4	6
EXPERTO5	9
VALORACIÓN MEDIA	7.8

Tabla 10.9
Puntuación subjetiva
para la evaluación de
validez de contenido.

En este caso, en la matriz de correlaciones no existe ninguna correlación con valor superior a 0.5, salvo las correlaciones de cada pregunta consigo misma (correlación = 1.0). Además, el número de encuestados ha sido de 37 (no llegando al mínimo de 50). Al analizarse 12 métodos en 4 dominios, el número de variables es igual a 48 variables, por lo que habría sido necesario un mínimo de 240 encuestados. Este último punto presenta muchas dudas, puesto que no hay estudios previos que puedan explicitar el número de variables, ya que podría considerarse como número de variables, el número de métodos evaluados, o bien, el número de dominios. En cualquier caso, al no contar con datos experimentales suficientes, no es posible decantarse por ninguna opción.

10.8.3 Validez de contenido

La validez de contenido de un cuestionario indica que las preguntas que lo componen forman una muestra adecuada del ámbito que se está evaluando. Este tipo de validación es la mínima garantía de validez que debe tener cuestionario. Se suele determinar a partir del análisis racional de las preguntas mediante un juicio de expertos en la materia. En este caso, se ha solicitado a 5 expertos en la materia (todos ellos, profesores universitarios con publicaciones en el ámbito del procesamiento de imágenes) que puntúen (entre 0 y 10) el grado de adecuación de los dos cuestionarios para evaluar diferentes métodos de extracción de contornos con imágenes con sombras tanto para personas con un grado de VISIÓN ESTÁNDAR como con individuos con BAJA VISIÓN. En la Tab. 10.9 se muestran las valoraciones proporcionadas por los expertos. Se ha obtenido una puntuación media de 7.8 sobre 10. Al obtener un valor superior a 5, se puede considerar que una mesa de expertos ha validado positivamente los cuestionarios.

Todos los expertos mostraron reticencias ante las preguntas de mapeo espacial y temporal, debido a que al repetirse la misma imagen con distintos algoritmos de contornos, cabe la posibilidad de que desde que por primera vez el usuario sea capaz de determinar la posición del objeto que se está buscando, siempre seleccione con el ratón la misma posición, aunque realmente no lo haya visto para dicho método en particular. Por otra parte, todos ellos mostraron su interés por dicho mecanismo, ya que les parecía novedoso y, con las salvedades anteriormente indicadas, altamente discriminante, ya que permite al encuestador establecer la veracidad de las respuestas de los encuestados.

10.9 FIABILIDAD DE LA ENCUESTA

COMO se ha comentado con anterioridad, la fiabilidad de las encuestas se puede medir desde dos enfoques distintos:

FIABILIDAD EXTERNA O TEMPORAL También llamada *Repetibilidad externa*. Evalúa si al aplicar los cuestionarios en distintos instantes temporales, bajo las mismas condiciones y sobre los mismos encuestados, se obtienen los mismos resultados. Para poder repetir los cuestionarios, estos deben ser cortos, concisos y directos, que permitan ser contestados en poco tiempo.

FIABILIDAD INTERNA Determina si diferentes encuestados, con similares características, proporcionan resultados análogos para una misma aplicación de la encuesta. La *Fiabilidad interna* aumenta al incrementar el número de preguntas y al realizar varias preguntas sobre el mismo aspecto, ya que suele fijar el interés de los encuestados y tiende a descartar respuestas *al azar*.

10.9.1 *Fiabilidad externa*

En esta Tesis Doctoral, no se está proponiendo la encuesta como escala universal, fija e intemporal de medida de la calidad de los métodos de extracción de contornos. Esta encuesta *ad hoc* pretende obtener una evaluación puntual de la calidad de los métodos de extracción de contornos con aplicación para usuarios con BAJA VISIÓN. De hecho, los cuestionarios no se han diseñado con el objetivo de *Fiabilidad temporal* en mente, sino con el objetivo de obtener una medida elevada de *Consistencia interna*, por lo que no se ha primado el tener un número reducido de preguntas (lo cual hubiese facilitado una nueva aplicación). En particular, sobre el conjunto de encuestados que finalizaron los dos cuestionarios, se realizó un sondeo para conocer su disponibilidad para rehacer de nuevo la encuesta, obteniendo respuestas negativas en este sentido, alegando su excesiva longitud. En dicho sentido, por tanto, se ha descartado su análisis. Sin embargo, la no realización de este análisis no afecta a la fiabilidad real de los resultados obtenidos en este estudio, sino a la previsión que se pudiese hacer para futuros estudios.

10.9.2 *Fiabilidad interna*

La *Fiabilidad interna* de una encuesta se puede obtener utilizando dos parámetros que permiten su medida: teniendo en cuenta la *Consistencia interna* de los cuestionarios y a partir del nivel del *Grado de acuerdo* entre los distintos encuestados.

Consistencia interna

Utilizando el estadístico α -Cronbach se ha podido determinar el nivel de consistencia interna de los cuestionarios en general y de cada uno de los métodos en particular. En la Tab. 10.10 se muestran los valores de dicho estadístico, tanto para cada uno de los dos cuestionarios, para

CUESTIONARIO	MÉTODO	α -Cronbach
CUESTIONARIO 1		0.913
	SOBEL-80	0.695
	CANNY-80	0.531
	LoG-95	0.729
	LIPCANNY-80	0.673
	LIPSOBEL-80	0.586
	LIPLoG-95	0.676
CUESTIONARIO 2		0.924
	SOBEL-90	0.635
	CANNY-90	0.651
	LoG-120	0.630
	LIPCANNY-90	0.734
	LIPSOBEL-90	0.795
	LIPLoG-120	0.767
ENCUESTA COMPLETA		0.956

Tabla 10.10
Cálculo del α -Cronbach de los cuestionarios y de cada método.

la encuesta completa, como para cada uno de los métodos evaluados en cada cuestionario.

Como se puede observar en la Tab. 10.10, tanto el α -Cronbach de cada cuestionario por separado como el α -Cronbach de la encuesta completa es muy cercano a 1.0, lo que quiere decir que existe una alta consistencia en las respuestas debido a que la varianza entre las mismas es muy baja con respecto al número total de preguntas. Por lo que se puede deducir que las preguntas no se han contestado al azar y por tanto, el efecto de difusión de la muestra de las respuestas proporcionadas por los encuestados es baja. Esto significa que los cuestionarios, en su globalidad, presentan una alta estructuración en sus preguntas, que es concordante con la baja variabilidad en las respuestas proporcionadas por los encuestados, lo cual imprime un alto grado de consistencia a la encuesta.

También se puede observar en la Tab. 10.10 que el α -Cronbach de cada método es menor, aunque superior a 0.5, lo cual implica que las respuestas de los encuestados centrados en cada método en particular presentan un poco de más variabilidad, lo que concuerda con las diferentes opiniones y calificaciones subjetivas de cada usuario. Sin embargo, aunque es algo mayor que en el caso de los cuestionarios o de la encuesta completa, el grado de consistencia de cada método también muestra que los encuestados no han contestado a las preguntas de forma aleatoria, sino que la variabilidad de las mismas está en un rango acotado y aceptable.

Es importante destacar el hecho de que el uso de preguntas que no son de puntuación (preguntas de selección y preguntas de mapeado temporal y espacial) modifica mucho el valor de α -Cronbach como se muestra en la Tab. 10.11, en la que se comparan el α -Cronbach de todas las preguntas que involucren a un determinado método y el α -Cronbach

Tabla 10.11
Comparación del α -Cronbach de cada método teniendo en cuenta todas las preguntas (TODAS) y solamente aquellas de puntuación (SÓLO PUNT.).

MÉTODO	α -Cronbach	
	TODAS	SÓLO PUNT.
SOBEL-80	0.695	0.742
CANNY-80	0.531	0.771
LoG-95	0.729	0.726
LIPSOBEL-80	0.586	0.755
LIPCANNY-80	0.673	0.800
LIPLoG-95	0.676	0.689
SOBEL-90	0.635	0.715
CANNY-90	0.651	0.700
LoG-120	0.630	0.668
LIPSOBEL-90	0.795	0.852
LIPCANNY-90	0.734	0.851
LIPLoG-120	0.767	0.752

calculado tomando únicamente las preguntas de puntuación relativas a dicho método. En este caso, el α -Cronbach tomando únicamente las preguntas de puntuación eleva su valor con respecto al cálculo utilizando todas las preguntas, ya que las preguntas de selección son dicotómicas y por tanto, se traducen a los valores extremos de la escala *Lickert* común (0 ó 10), con lo que en el caso de que existan algunas discrepancias entre los encuestados, la varianza está sobreponderada hacia los valores extremos, siendo mucho mayor que en el caso de las preguntas de puntuación.

En la Fig. 10.5 se muestran las medias de las valoraciones de cada encuestado, agrupadas por métodos. Se puede observar que cada método presenta un alto agrupamiento (baja difusión) de las puntuaciones, lo que favorece que el cálculo del α -Cronbach sea elevado.

Grado de acuerdo

Se ha realizado el estudio del *Grado de acuerdo* mediante el cálculo del estadístico ICC que permite determinar el promedio de las correlaciones entre todas las posibles ordenaciones de los pares de observaciones disponibles y, por lo tanto, evita el problema de la dependencia del orden del coeficiente de correlación. Además, permite extender su uso a más de dos observaciones por sujeto. Precisamente esos dos motivos hacen que este coeficiente sea más robusto que el coeficiente *r* de *Pearson*.

Siguiendo la metodología propuesta por *Shrout y Fleiss* [99], se ha calculado el ICC a partir de un análisis de varianza de doble vía utilizando el valor medio de las respuestas de los encuestados como unidad básica de comparación. Por tanto, según la nomenclatura de *Shrout y Fleiss* es un ICC(3,k). En la Tab. 10.12 se presentan los cálculos de ICC para los dos cuestionarios, tomando en consideración todos los encuestados (columna TODOS). También se calculan tanto el ICC sobre las respuestas de los encuestados con VISIÓN ESTÁNDAR (columna VE)

ICC: Del inglés, INTRACLASS CORRELATION COEFFICIENT. Coeficiente de Correlación Intraclass.

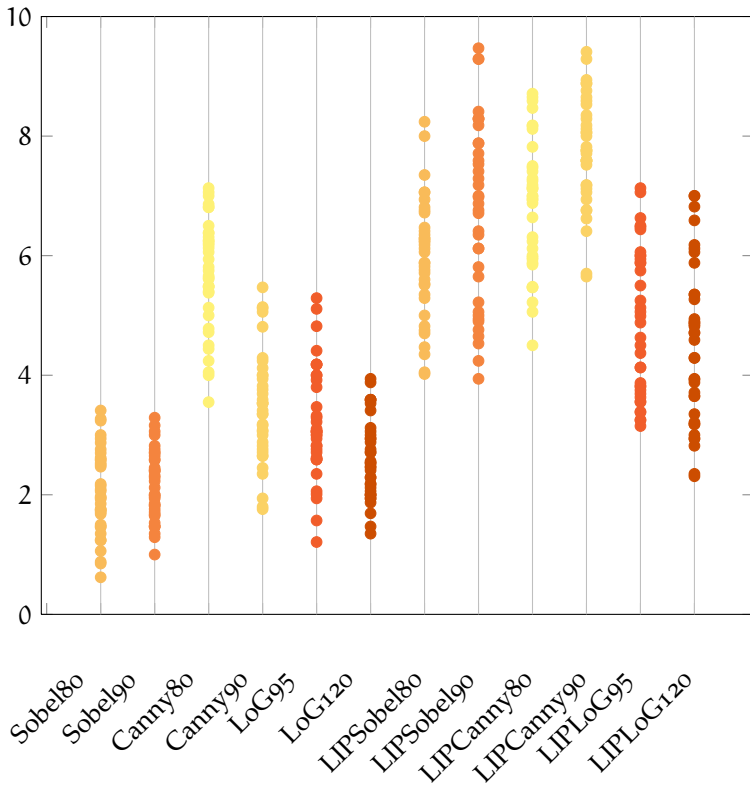


Figura 10.5 Gráfica de valoración media de cada usuario por método.

CUESTIONARIO	ICC		
	TODOS	VE	BV
CUESTIONARIO 1	0.992	0.990	0.950
CUESTIONARIO 2	0.994	0.994	0.956
ENCUESTA COMPLETA	0.993	0.992	0.962

Tabla 10.12 Cálculo del ICC de los cuestionarios según tipo de usuario. BV: Individuos con Baja Visión. VE: Individuos con visión estándar.

como el ICC de las respuestas de aquellos encuestados con problemas de BAJA VISIÓN (columna BV). Todos los valores de ICC calculados muestran que existe un alto grado de correlación entre las respuestas de los encuestados en cada uno de los cuestionarios. También se puede deducir que tanto las respuestas de los encuestados con BAJA VISIÓN como las respuestas de los encuestados con un grado de VISIÓN ESTÁNDAR presentan un grado de acuerdo muy elevado.

10.10 CONCLUSIONES

En este capítulo se han presentado los resultados de la encuesta de evaluación de algoritmos de extracción de contornos, constituida por dos cuestionarios diseñados *ex profeso* para esta Tesis Doctoral. También se ha mostrado el proceso completo desde el pilotaje de las primeras versiones de los cuestionarios hasta la validación de la encuesta,

pasando por el proceso de selección de individuos de la población o el uso de recordatorios durante la fase de contestación de la encuesta, entre otros aspectos relevantes.

Las conclusiones obtenidas se pueden subdividir en dos grandes apartados:

- Estudio, diseño y proceso de respuesta de los cuestionarios.
- Análisis de los resultados de las respuestas de los encuestados.

10.10.1 Conclusiones sobre el proceso de respuesta

El propio proceso de análisis y diseño de estos cuestionarios tiene suficiente entidad como para contemplarlo dentro de los resultados de esta investigación. Esta parte engloba todos los aspectos previos al cierre de los cuestionarios, último paso antes de la obtención de los resultados finales. En esta investigación se han obtenido datos relevantes de los siguientes aspectos:

PILOTAJE PREVIO Realizado con 5 usuarios y sirvió para obtener gran parte de las recomendaciones de diseño de los cuestionarios con contenidos visuales.

PILOTAJE FINAL Aplicado sobre 4 voluntarios, permitió detectar y eliminar prácticamente la totalidad de los errores de los cuestionarios finales.

VISIBILIDAD DE LA ENCUESTA Inicialmente, se estableció que la encuesta debería permanecer abierta y disponible durante 12 semanas, aunque debido a la baja TASA DE RESPUESTA se amplió a 16 semanas.

SELECCIÓN DE LA POBLACIÓN Se escogió el método de “bola de nieve” a partir de una población inicial de 26 individuos, obteniendo una población definitiva de 70 individuos, lo que ha significado un aumento en la TASA DE MODIFICACIÓN DE LA POBLACIÓN MUES-TRAL de un 269.2 %.

EFFECTO DE LOS RECORDATORIOS Se estudió el efecto de las cartas de recordatorio, encontrándose que provoca un efecto positivo tanto en el número de usuarios registrados en el sistema de cuestionarios como en el número de encuestas completadas en el mismo. Para esta encuesta se realizaron 2 recordatorios, consiguiéndose con el primero de ellos un aumento de 8 usuarios en tan sólo tres días (lo que representa un aumento de la población del 116.67 %) y de 17 usuarios en quince días (significando un aumento del 135.42 %). Mientras que con el segundo recordatorio, se aumentó en 3 usuarios durante los 3 primeros días (aumento del 104.62 %) y a 5 tras 15 días (para un aumento del 107.69 %). De manera similar, tras el primer recordatorio, el número de encuestas finalizadas aumentó de 12 a 27 en quince días (aumento del 225 %) y tras el segundo recordatorio, el número de encuestas completadas subió de 27 a 37 en los últimos 15 días de vigencia de la encuesta (aumento del 137.04 %).

TASA DE RESPUESTA Se ha calculado este valor, como medida de la cantidad de usuarios que han respondido a la encuesta en relación con la población completa. En esta investigación, establecer el tamaño de la población completa es muy complejo y se ha conseguido estimar que la carta de presentación llegó, al menos, a 190 individuos. Tomando estos datos, se ha obtenido un valor de TASA DE RESPUESTA con respecto al número encuestas finalizadas, del 19.5 % de la población completa estimada y de un 52.9 % con respecto a aquellos usuarios que mostraron interés en la misma y se registraron en la aplicación Web. Estas tasas, aún siendo bajas, están dentro del rango habitual de las encuestas auto-administradas por Internet y sólo se debe tener en cuenta la TASA DE RESPUESTA ADAPTADA, que es la que tiene como referencia el número de usuarios registrados en la aplicación, puesto que la población completa es una mera estimación.

De todo lo expuesto a lo largo del capítulo, se pueden sacar las siguientes conclusiones:

- Para desarrollar una encuesta con garantías es necesario realizar el diseño pasando por uno o más filtros de calidad, obtenidos a partir de diversos "pilotajes", que deben ser realizados por un número pequeño y controlado de usuarios, que no deberían involucrarse en las etapas posteriores.
- Para encuestas auto-administradas por Internet, como la de esta investigación, el periodo que debe estar disponible la encuesta para que la contesten los encuestados debe ser como mínimo de 4 meses.
- Para la selección de la población objetivo es preferible utilizar métodos que permitan su selección previa bajo los estándares que se determinen. Sin embargo, si lo anterior no es posible, el método "bola de nieve" es viable aunque hay que prestar un especial cuidado a la selección de la población inicial.
- La población inicial en un método "bola de nieve" debe garantizar una TASA DE DIFUSIÓN INICIAL muy alta. En caso de que sean necesarios varios estratos (individuos con características diferentes) será imprescindible tener suficientes individuos de cada uno de dichos estratos en la población inicial.
- La encuesta asociada a esta investigación presenta una TASA DE RESPUESTA dentro del rango habitual, ya que este tipo de encuestas suele mostrar un bajo nivel de respuestas (entorno al 20-25 % de la población completa), aunque si se tiene en cuenta únicamente aquellos encuestados que han mostrado cierto grado de interés (por ejemplo, registrándose en la aplicación) la TASA DE RESPUESTA ADAPTADA suele subir hasta el 50 %.

De manera adicional, algunos encuestados al finalizar los dos cuestionarios proporcionaron su opinión personal sobre el proceso de respuesta de los mismos. De estos se obtienen unas conclusiones de tipo cualitativo:

- Los dos cuestionarios son excesivamente largos, tanto en número de preguntas como en cantidad de imágenes de respuesta.
- Las preguntas de mapeado espacial y temporal tienen una utilidad relativa, ya que al utilizar la misma imagen original y preguntar siempre por el mismo elemento, la calidad real de la respuesta está limitada por la pregunta en la que se detecte dicho elemento, ya que, a partir de esa pregunta el encuestado sabe dónde se encuentra el elemento que se desea buscar.
- El uso de dos cuestionarios con exactamente las mismas preguntas (aunque varíen las imágenes de respuesta) desconcierta a los encuestados.

Finalmente, algunos de los encuestados propusieron algunas soluciones a los problemas que ellos mismos detectaron. La solución más importante es la relativa a las preguntas de mapeado espacial y temporal, en las que propusieron el uso de imágenes originales similares en las que se cambiase el objeto que se desea detectar y la posición de los objetos principales en la escena. Esto podría garantizar un cierto nivel de uniformidad y, si se garantiza que la selección del método y de la imagen original es totalmente aleatorio, no favorecería a ningún método frente a otro.

10.10.2 Conclusiones sobre los resultados

Tras la contestación de la encuesta por parte de los usuarios, lo primero que se ha realizado es analizar *a posteriori* la validez de cada individuo en función de las respuestas proporcionadas. De esta forma, se pueden eliminar OUTLIERS y se puede garantizar que cualquier deducción que se haga sobre los datos es representativa dentro de un rango de funcionamiento normal. Para ello, se ha calculado la media de las respuestas de cada encuestado para cada método y se ha analizado el conjunto de dichos datos para comprobar que siguen una distribución normal. Mediante el test de *Shapiro-Wilk* se ha comprobado, con un nivel de significación del 95 %, que la respuesta media de cada encuestado sigue una distribución gaussiana. Gracias a lo cual, se ha aplicado el método de *Grubb* para la detección de OUTLIERS, no detectándose ninguno. Por ello, se han podido utilizar las respuestas de todos los encuestados para el análisis y deducción de los resultados finales. Utilizando estos resultados, con carácter general, se ha podido determinar que:

MÉTODOS LIP ESTÁN MEJOR VALORADOS Todos los métodos LIP estudiados obtienen mejores calificaciones subjetivas que los respectivos métodos en su versión No LIP.

MÉTODOS LIP APRUEBAN La mayoría de métodos LIP obtienen valoraciones superiores al 5 y los que no aprueban, se quedan cerca de superar dicho valor.

ALTA CONSISTENCIA INTERNA Todos los métodos, tanto los LIP como los No LIP, presentan una baja variabilidad, por lo que el grado de consistencia interno, medido a través del estadístico α -Cronbach, es elevado.

BAJA DIFUSIÓN EN CADA MÉTODO Los encuestados muestran un grado de difusión bastante bajo en sus respuestas, sobre todo si se agrupan las respuestas por método. Esto demuestra que existe una correlación entre clases muy elevada, corroborado mediante el cálculo del estadístico $ICC(3,k)$.

Utilizando el perfil de cada encuestado se ha podido establecer el grado de visión de cada usuario a partir de la puntuación proporcionada por lo mismos a una serie de problemas de visión específicos. Este análisis permitió catalogar los 37 encuestados en dos tipos diferenciados: 32 individuos que poseen un grado de VISIÓN ESTÁNDAR, y 5 usuarios que presentan problemas de BAJA VISIÓN. A partir de esta diferenciación entre usuarios se han podido obtener otras conclusiones, de carácter más específico, que relacionan las respuestas obtenidas con el grado de visión de cada uno de estos dos grupos de encuestados. A continuación, se enuncian las conclusiones más importantes obtenidas en este aspecto concreto.

LOS DOS GRUPOS MUESTRAN LA MISMA TENDENCIA Tanto los usuarios con VISIÓN ESTÁNDAR como los encuestados con BAJA VISIÓN muestran una tendencia similar en sus respuestas para todos los métodos evaluados, como lo demuestra el alto grado de correlación obtenido entre las respuestas de los dos grupos. Esto quiere decir que cuando un grupo puntúa de manera elevada a un método en una determinada pregunta, el otro grupo también lo puntuará de manera elevada; y viceversa.

LAS PUNTUACIONES SON PROPORCIONALES Se ha demostrado que las respuestas de los encuestados de los dos grupos están relacionadas entre sí de manera proporcional, mostrando que las respuestas medias de los encuestados con VISIÓN ESTÁNDAR son superiores entre un 10 % y un 45 % a las respuestas de los usuarios con BAJA VISIÓN, dependiendo del método.

MÉTODOS LIP CON PROPORCIÓN DE MENOR VARIABILIDAD Los encuestados de ambos grupos puntúan de manera muy similar los métodos LIP, lo que hace que la relación entre las respuestas de ambos grupos de encuestados muestren una variación en la proporción menor que para los métodos No LIP.

RELACIÓN LINEAL ENTRE LOS GRUPOS Se ha calculado la recta de regresión con dos grados de libertad entre las respuestas medias de cada grupo de usuario para cada uno de los métodos evaluados, obteniéndose un ajuste con un error muy reducido. Dicho ajuste permite asegurar que las respuestas de los encuestados con VISIÓN ESTÁNDAR y las de los que sufren de BAJA VISIÓN son idénticas, sólo que las del primer grupo están sobreestimadas en aproximadamente un punto.

Todas estas afirmaciones y conclusiones expuestas anteriormente se han visto reforzadas por la validación de la encuesta mediante la puntuación proporcionada por los expertos en la validación de contenido,

y por el estudio de fiabilidad de la encuesta, mediante una alta consistencia interna obtenida a través del estadístico α -Cronbach y una elevada correlación intra-clases calculada mediante ICC(3,k). Estos cálculos han permitido demostrar que la encuesta que se ha construido y utilizado para la evaluación de la calidad de los contornos, sirve para medir fielmente este objetivo. Todo ello, a pesar de que la encuesta no se ha diseñado como métrica universal, sino que es una herramienta *ex profeso* para esta Tesis Doctoral.

10.10.3 Conclusiones de la Parte III

Este capítulo cierra la Parte III, *Resultados Experimentales*, que demuestra la utilidad del planteamiento teórico expuesto en los diversos capítulos de la Parte II.

Esta Parte III se ha dividido en 2 capítulos. El Capítulo 9 ha mostrado el funcionamiento de la operación *Convolución*-LIP con distintos métodos de extracción de contornos, comparando tanto visualmente como en tiempos de computación, el método propuesto con otras aproximaciones, tanto LIP como No LIP.

En el Capítulo 10 se ha mostrado todo el proceso de captación de las opiniones de los usuarios a través de cuestionarios. Gracias a estos cuestionarios, se ha confirmado que los encuestados prefieren las imágenes de contornos extraídas utilizando los métodos LIP frente a los que no lo utilizan. Para finalizar, se ha podido determinar que tanto los usuarios con un grado de VISIÓN ESTÁNDAR como aquellos que sufren de algún tipo de problema de BAJA VISIÓN muestran un nivel de correlación muy alto entre sus respuestas, lo cual, permite afirmar que tanto unos como otros presentan un mismo modelo conceptual de calidad de contornos.

Parte IV

DISCUSIÓN Y CONCLUSIONES

DISCUSIÓN

ÍNDICE

11.1	Discusión sobre el operador <i>Convolución-LIP</i>	220
11.1.1	Doble implementación	220
11.1.2	Fácil adaptación a LIP	221
11.1.3	Mejor comportamiento en zonas con baja iluminación	221
11.1.4	Ajuste robusto de parámetros	221
11.1.5	Comportamiento unificado sin pasos adicionales	222
11.2	Discusión sobre las encuestas de evaluación de contornos	222
11.2.1	Métodos LIP mejor valorados	223
11.2.2	Métodos LIP aprueban	223
11.2.3	Encuestados puntúan de manera similar	224
11.2.4	Respuestas proporcionales entre grupos	224
11.3	Conclusiones	224

En este capítulo se analizan y discuten los resultados obtenidos a lo largo de la presente Tesis Doctoral. Se muestran las principales aportaciones científicas que esta investigación proporciona, haciendo hincapié en las fortalezas que presenta tanto el nuevo operador *Convolución-LIP* propuesto como el mecanismo de evaluación de la calidad de los algoritmos de extracción de contornos mediante encuestas de opinión. Primero, se discuten los aspectos relativos al operador *Convolución-LIP* (cuyo desarrollo representa una de las mayores contribuciones de este trabajo), donde se analizan y discuten los resultados de los diferentes experimentos realizados.

Tras ello, la discusión se centra en el proceso de evaluación de los algoritmos propuestos mediante las encuestas de opinión. Se estudian los resultados obtenidos tanto por los diversos algoritmos implementados como agrupados según el nivel de visión de los encuestados.

11.1 DISCUSIÓN SOBRE EL OPERADOR *convolución-LIP*

Una de las principales aportaciones científicas es el desarrollo del operador *Convolución-LIP*, que es la adaptación al paradigma LIP del operador tradicional *convolución* 2D con filtros separables. El desarrollo matemático de este operador se ha expuesto en el Cap. 6. A modo de resumen, se puede afirmar que este nuevo operador presenta los siguientes puntos innovadores.

Tabla 11.1
Listado de características principales del nuevo operador *Convolución-LIP*.

- A. Existen dos posibles implementaciones: DGLIP-CONV, que utiliza la aproximación “directa” y FGLIP-CONV, que hace uso de operaciones computacionalmente más sencillas y que, por tanto, es más rápida y computacionalmente más ligera que la anterior.
- B. Permite adaptar al paradigma LIP todos los algoritmos de PROCESAMIENTO DE IMÁGENES que utilicen la convolución 2D con filtros separables.
- C. Las versiones LIP de los algoritmos implementados con el operador *Convolución-LIP*, presentan un mejor comportamiento en imágenes con iluminación no uniforme.
- D. El ajuste de los algoritmos LIP es más robusto y menos sensible a pequeños cambios que para los mismos algoritmos en sus versiones No LIP.
- E. Para obtener resultados similares en imágenes con iluminación no uniforme, los algoritmos No LIP deben hacer uso de métodos adicionales, que hacen que el proceso completo sea más lento que utilizando únicamente el algoritmo en su versión LIP.

A continuación, se refrendan todas las afirmaciones de la Tab. 11.1.

11.1.1 Doble implementación

El punto A de la Tab. 11.1 queda demostrado en el Cap. 6, ya que en él se ha establecido matemáticamente el operador *Convolución-LIP* y se ha mostrado la existencia de dos implementaciones alternativas diferentes para dicho operador. La implementación DGLIP-CONV utiliza la aproximación “directa” del paradigma LIP, construyendo la formulación a partir de multiplicaciones y exponenciaciones. Se puede considerar esta implementación como una generalización del método propuesto por Deng y Pinoli [29], ya que al utilizar los filtros con los que esos autores construyeron su adaptación, se obtiene exactamente la misma formulación en ambos casos.

También, en dicho Cap. 6, se formula otra implementación, FGLIP-CONV, que utiliza sumas y multiplicaciones, siendo éstas menos costosas a nivel computacional que las utilizadas en DGLIP-CONV.

11.1.2 *Fácil adaptación a LIP*

En el Cap. 9, se exponen diversos experimentos en los que se ha mostrado cómo, utilizando el operador *Convolución*-LIP en lugar del operador tradicional *convolución*, se puede adaptar de manera sencilla prácticamente cualquier algoritmo de PROCESAMIENTO DE IMÁGENES al paradigma LIP. La implementación, en dicho capítulo, de cinco algoritmos diferentes de distintos ámbitos corrobora el punto B de la Tab. 11.1. Se ha afirmado que la adaptación a LIP es fácil y sencilla, lo cual se demuestra en dicho capítulo, ya que únicamente se requiere la obtención de los dos filtros unidimensionales que forman el filtro bidimensional del método No LIP original.

Por el contrario, la propuesta de Deng y Pinoli [29] para portar algoritmos a LIP es extremadamente compleja y tediosa, ya que requiere obtener, usualmente *off-line*, la formulación del filtro bajo el paradigma LIP y programarlo. Todos estos pasos son imprescindibles, incluso para aspectos menores como puede ser ampliar o reducir el entorno de procesamiento. Sin embargo, para el operador *Convolución*-LIP basta con modificar el tamaño del filtro que se desee aplicar.

11.1.3 *Mejor comportamiento en zonas con baja iluminación*

La justificación del punto C de la Tab. 11.1 no es exclusiva de los algoritmos implementados con el operador propuesto *Convolución*-LIP, sino que es un efecto beneficioso que obtienen todos los algoritmos implementados bajo LIP. En general, todos estos métodos muestran un rendimiento más homogéneo en todas las zonas de la imagen, independientemente de la iluminación local de cada una de ellas. Se puede observar este comportamiento en las Fig. 9.3 y 9.4 en la pág. 160 del Cap. 9. La primera de ellas muestra la imagen magnitud del gradiente del algoritmo estándar *Sobel*, mientras que la segunda muestra la respuesta del algoritmo *LIP-Sobel* ante la misma imagen de entrada. En la primera imagen, los contornos de los pimientos de la zona izquierda (que se encuentran en la zona de sombras) se muestran débilmente y no tienen una intensidad constante en toda la imagen. Por el contrario, en el segundo caso, los bordes aparecen con una intensidad fuerte y constante en toda la extensión de la imagen. Se observa el mismo comportamiento comparando la Fig. 9.12, en la pág. 170, que proporciona la salida umbralizada del algoritmo *Canny*, con la Fig. 9.14, pág. 170, que muestra la salida del algoritmo *LIP-Canny*. En la imagen generada por *Canny*, no se obtienen los contornos de las ventanas que se encuentran bajo los soporales y sin embargo, estos aparecen en la salida del método *LIP-Canny*.

11.1.4 *Ajuste robusto de parámetros*

Un comportamiento deseable en todos los algoritmos es que el ajuste de los parámetros de funcionamiento no sea excesivamente complejo. Se ha podido demostrar empíricamente que el ajuste de los parámetros de funcionamiento de los métodos que se implementan utilizando *Convolución*-LIP es más robusto y más flexible que los mismos algorit-

mos en su versión No LIP. En el Cap. 9, se puede observar cómo el ajuste de los umbrales inferior y superior de *LIP-Canny* no necesita ser tan preciso como en el caso de los algoritmos *Canny* o para el método híbrido *Filtrado Homomórfico* seguido de *Canny*. En la Fig. 9.12, situada en la pág. 170, se muestra la salida del algoritmo *Canny* con umbral superior, $th_{high} = 0.125$, y umbral inferior, $th_{low} = 0.05$. La Fig. 9.13, pág. 170, representa la salida del método híbrido *Filtrado Homomórfico* seguido de *Canny*, con unos umbrales, $th_{high} = 0.05$ y $th_{low} = 0.01$. Finalmente, el mapa de contornos generados por *LIP-Canny* se ha presentado en la Fig. 9.14 (pág. 170). Para este último método, los umbrales seleccionados han sido $th_{high} = 0.5$ y $th_{low} = 0.04$. Como se puede determinar por observación directa de dichos umbrales, el algoritmo *LIP-Canny* tiene un rango de selección de umbrales mucho mayor, aspecto éste que proviene de ser un método implementado en el paradigma LIP. Esto es debido a que LIP provoca una normalización implícita de los rangos, además de reforzar más los saltos de contraste, con lo cual, se consigue que, para el caso particular de *LIP-Canny*, los picos de la magnitud del gradiente sean más elevados y estén mucho más reforzados que para *Canny*. Este comportamiento, aunque particularizado en una imagen, se puede extender a más métodos y ejemplos, y por tanto, demuestra lo afirmado en el punto D de la Tab. 11.1.

11.1.5 Comportamiento unificado sin pasos adicionales

Existen otras aproximaciones tradicionales que han permitido solventar o, al menos, limitar el efecto de la iluminación no uniforme en imágenes naturales. En el Cap. 9, se ha utilizado un mecanismo similar al indicado, la aplicación del *Filtrado Homomórfico* como paso previo a la obtención de contornos mediante *Canny*. El *Filtrado Homomórfico*, descrito en la Sec. 2.1.1, elimina la componente de *iluminación* de la imagen original mediante la aplicación del logaritmo neperiano y la resta del filtrado paso baja de la imagen original, generando una imagen que contiene únicamente la *reflectancia*, por lo que el método *Canny* extrae los contornos de los objetos presentes en esta última. Por el contrario, los métodos basados en LIP no requieren estos pasos, ya que este paradigma presenta un comportamiento logarítmico y siempre que se aplique un filtrado con un filtro cuyas componentes sumen 0, se garantiza que las componentes se anulan entre sí y el efecto sobre la imagen es equivalente al de una resta. El punto E de la Tab. 11.1 se demuestra mediante este efecto matemático.

11.2 DISCUSIÓN SOBRE LAS ENCUESTAS DE EVALUACIÓN DE CONTORNOS

Las encuestas para la evaluación de la calidad de los diferentes algoritmos de extracción de bordes en imágenes proponen una interesante discusión científica. Debido a que el uso de encuestas en el ámbito del PROCESAMIENTO DE IMÁGENES Y VÍDEO no está muy extendido, la cantidad de matices sobre los que se puede discutir es ingente. Sin embargo, para poder ser prácticos, es necesario concretar a unos pocos aspectos principales sobre los que argumentar.

- A. Los métodos implementados bajo LIP obtienen mejor valoración que las respectivas versiones No LIP.
- B. La mayoría de los métodos implementados en LIP aprueban en su valoración.
- C. Los encuestados con VISIÓN ESTÁNDAR y los usuarios con BAJA VISIÓN puntúan de manera similar los diferentes métodos de extracción de contornos.
- D. Las puntuaciones de los usuarios con VISIÓN ESTÁNDAR es proporcional a las proporcionadas por los encuestados con BAJA VISIÓN.

Tabla 11.2
Listado de características principales de la evaluación de los algoritmos de extracción de contornos mediante encuestas de opinión.

En los siguientes apartados se muestran las discusiones que permiten la demostración de cada uno de los puntos de la Tab. 11.2.

11.2.1 Métodos LIP mejor valorados

Los resultados expuestos en el Cap. 10 muestran que los usuarios proporcionan puntuaciones más altas a los métodos *LIP-Sobel*, *LIP-LoG* y *LIP-Canny*, que a los métodos *Sobel*, *LoG* y *Canny*, lo cual demuestra el punto A de la Tab. 11.2. Con ayuda de la Tab. 10.4 y la Fig. 10.2 (pág. 201), que permiten comparar todos los métodos evaluados, se puede observar que, en todos los casos, las puntuaciones de los métodos implementados bajo LIP son superiores a las puntuaciones de los mismos métodos en su versión clásica (No LIP). El único caso que merece la pena destacar como alta calidad de entre los algoritmos No LIP es el de *Canny-80*, es decir el algoritmo clásico de *Canny* con el 80 % de los píxeles de la imagen, como mínimo, marcados como no pertenecientes a ningún contorno. Dicho algoritmo obtiene una valoración de 5.68 (sobre 10), que se acerca en su evaluación a *LIPCanny-80* (que obtiene un 6.87).

11.2.2 Métodos LIP aprueban

De nada serviría que los métodos LIP obtengan una mejor valoración que los métodos No LIP equivalentes, si los primeros no consiguiesen obtener una puntuación suficiente para “aprobar”. En este caso, los resultados expuestos en la Tab. 10.4 y la Fig. 10.2 (pág. 201) permiten demostrar el punto B de la Tab. 11.2, aunque hay que matizar ligeramente esta afirmación, ya que dos de los métodos no obtienen una valoración superior al 5, aunque es cercana a ésta. Se trata de los métodos *LIPLoG-95* y *LIPLoG-120*, que han sido valorados, respectivamente, con 4.53 y 4.17 sobre 10.

11.2.3 Encuestados puntúan de manera similar

Uno de los grandes interrogantes antes de realizar esta Tesis Doctoral era saber si los encuestados con un grado de VISIÓN ESTÁNDAR tenían el mismo concepto mental de calidad de contornos que los usuarios con BAJA VISIÓN. El cálculo del coeficiente de correlación r de Pearson entre los dos grupos de individuos permite determinar si existe algún tipo de relación entre ambos grupos. En este caso, la Tab. 10.6 muestra que la correlación entre los dos grupos ha sido de 0.9874, lo cual permite demostrar el punto **c** de la Tab. 11.2. Un valor de correlación tan elevado quiere decir que, prácticamente en la totalidad de casos, siempre que un usuario con VISIÓN ESTÁNDAR valora positivamente un mapa de contornos, un encuestado con BAJA VISIÓN también lo valora positivamente, y viceversa.

11.2.4 Respuestas proporcionales entre grupos

La consecuencia más directa del apartado anterior, consiste en determinar si existe una relación proporcional entre las respuestas de ambos grupos. Esto, que es lo que se afirma en el punto **d** de la Tab. 11.2, se demuestra en el Cap. 10. En particular, en la Tab. 10.7 se observa cómo los valores de las respuestas medias de los usuarios con VISIÓN ESTÁNDAR son superiores entre un 10% y un 45% respecto a los valores de las respuestas de los encuestados con BAJA VISIÓN. Estos valores sugieren la existencia de una relación proporcional entre las respuestas de los usuarios de los dos grupos. Para corroborar esta afirmación, se ha realizado la regresión lineal de las respuestas de ambos grupos de usuarios, determinándose que existe una relación lineal que sigue la ecuación de ajuste $Y = 0.9970 \cdot X + 0.9397$, donde Y es la respuesta de un usuario medio con VISIÓN ESTÁNDAR y X representa la respuesta de un encuestado tipo con BAJA VISIÓN. Este ajuste permite afirmar que existe una relación lineal entre las respuestas de ambos grupos, en la que los usuarios con BAJA VISIÓN proporcionan valoraciones casi un punto menor que las que indican los encuestados con VISIÓN ESTÁNDAR.

11.3 CONCLUSIONES

Los resultados obtenidos en los experimentos avalan los avances e innovaciones científicas que se han indicado a lo largo de la Tesis y que se han resumido en este capítulo. Se han dividido en dos grandes bloques las revisiones de las aportaciones de investigación, las relativas al nuevo operador propuesto, *Convolución-LIP*, y aquellas relacionadas con las encuestas de evaluación de los algoritmos de extracción de bordes.

Se han seleccionado cinco características innovadoras aparecidas gracias al desarrollo y uso del operador *Convolución-LIP*. Estas características se han considerado las más interesantes y relevantes para su discusión, lo cual no significa que sean las únicas que se pueden conseguir a partir de dicho operador.

En la sección sobre los resultados obtenidos del diseño y aplicación de

las encuestas de evaluación de la calidad de los métodos de extracción de contornos, se han resaltado y discutido cuatro de las características más relevantes de los experimentos descritos en el Cap. 10.

En el próximo capítulo se indicarán las conclusiones finales de la presente Tesis Doctoral, así como la descripción futuras vías de investigación que pudieran surgir a partir de este trabajo.

CONCLUSIONES

ÍNDICE

12.1	Sumario	228
12.2	Producción científica	232
12.3	Principales aportaciones científicas	233
12.3.1	Sobre el operador Convolución-LIP	233
12.3.2	Sobre las encuestas de evaluación de contornos	233
12.4	Investigaciones futuras	235

Este capítulo sirve como colofón del trabajo realizado a lo largo de la presente Tesis Doctoral. En él, se resume el trabajo realizado en cada capítulo. También, se indican las conclusiones y las principales aportaciones científicas de este trabajo, así como posibles ampliaciones futuras en esta línea de investigación.

El capítulo se estructura en cuatro secciones. En la primera de ellas, se realiza un sumario, con los aspectos más relevantes de cada capítulo, y que sirve, a su vez, para reflejar todo el trabajo realizado.

En la segunda sección, se enuncian las publicaciones científicas que se han generado a partir del trabajo realizado en esta Tesis.

Tras ello, la siguiente sección sirve para exponer las conclusiones alcanzadas, tanto en relación con los objetivos propuestos inicialmente, como aquellos aspectos que no se habían contemplado en un principio, pero que se han logrado a lo largo de este trabajo. En este apartado, también se hace un especial hincapié en indicar las aportaciones científicas de esta Tesis.

Para finalizar, se exponen nuevas vías de investigación posibles que esta Tesis Doctoral abre para el futuro.

12.1 SUMARIO

El trabajo realizado en esta Tesis Doctoral es una contribución en el ámbito del desarrollo de algoritmos de extracción de contornos para escenas naturales, que contengan iluminación no uniforme en toda la imagen o sombras proyectadas. Este trabajo representa un avance en la evaluación subjetiva de la calidad percibida de los bordes extraídos mediante encuestas de opinión.

El trabajo realizado y las principales aportaciones de cada capítulo se resumen a continuación:

Capítulo 1

En él se ha realizado la introducción a la Tesis. Se han expuesto las motivaciones tanto sociales como técnicas que han sugerido el desarrollo de este trabajo de investigación. Se han descrito múltiples campos de aplicación, entre los cuales se ha escogido el de la ayuda a los pacientes con BAJA VISIÓN. Se ha fijado el marco de trabajo, los OBJETIVOS PRIMARIOS y SECUNDARIOS, y una propuesta que ha servido como punto de arranque de toda la investigación descrita en la memoria de esta Tesis Doctoral.

Capítulo 2

En dicho capítulo se ha realizado una revisión de diferentes métodos de PROCESAMIENTO DE IMÁGENES. Se ha expuesto en profundidad todos los fundamentos matemáticos del paradigma LIP, así como, las leyes que se cumplen al utilizarlo. Dentro del grupo de técnicas utilizadas para eliminar el efecto de la iluminación no uniforme, se ha descrito el *Filtrado Homomórfico*. Se ha realizado una revisión de las técnicas de extracción de contornos más habituales, centrándose en aquellos métodos que utilizan filtros de *convolución* para la obtención de los mapas de bordes.

Capítulo 3

Puesto que hay que evaluar la calidad de los mapas de contornos generados, en este capítulo se describe la importancia de los bordes en la visión y se hace un estudio de los métodos existentes para determinar la calidad en imágenes procesadas. Se exponen las dos grandes familias de mecanismos: los métodos OBJETIVOS y los SUBJETIVOS. Dentro de este segundo tipo de evaluación, se describen la métrica MOS diseñada por la UIT, la extensión sugerida por Durucan *et al.* [31] y la metodología propuesta por Heath *et al.* [47–50]. Esta última sirve de base para la creación de encuestas de opinión para la evaluación de la calidad de contornos y es la que se ha utilizado como punto de inicio para el desarrollo de los cuestionarios en esta Tesis. En este capítulo se explica por qué para las imágenes, en particular, se debe realizar la evaluación de la calidad por mecanismos SUBJETIVOS.

Capítulo 4

Al hacerse necesaria la evaluación mediante encuestas, se hace conveniente realizar un estudio de la metodología de investigación basada en cuestionarios. Al analizarla se ha encontrado que la mayoría de cuestionarios se definen para encuestas de tipo lingüístico en los ámbitos sociales o económicos, no habiendo prácticamente referencias validadas para la realización de encuestas en el campo del PROCESAMIENTO DE IMÁGENES.

Se han estudiado las fuentes de error más habituales, denominadas *sesgos*. Se ha podido reconocer un nuevo *sesgo*, del que no se tenía constancia que hubiera sido identificado previamente. Es el llamado *sesgo por finalidad*, que se produce cuando el encuestado no responde su opinión personal e independiente sino que lo hace condicionado por saber la finalidad de la pregunta, en particular, o de la encuesta, en general. Este tipo de *sesgo* es muy probable que aparezca en cuestionarios con preguntas con imágenes, por lo que, esta nueva fuente de error debe ser tenida en cuenta para la evaluación subjetiva dentro del campo del PROCESAMIENTO DE IMÁGENES.

También se incluye en este capítulo, una sección de ejemplos de preguntas erróneas con la explicación del motivo y un posible enunciado que lo corrige. Ésta, aunque de carácter menor, también es una aportación novedosa, ya que ayuda a identificación de errores comunes en la redacción de las preguntas de las encuestas. Para finalizar, se describen los mecanismos de validación de las encuestas, que garantizan (dentro de unos ciertos límites) que dicho cuestionario es un buen mecanismo de medida de la variable estadística que se desea estimar.

Capítulo 5

Debido a que en el Capítulo 1 se indicó que uno de los ámbitos de actuación es el BAJA VISIÓN, es necesario estudiar los cuestionarios existentes para dicha población, para conocer los aspectos principales que otros investigadores han estimado que afectan a dicho grupo de individuos. En este capítulo se realiza una revisión de 14 cuestionarios relacionados con la BAJA VISIÓN. Todos ellos son de tipo lingüístico y, en general, se han diseñado para la identificación de la BAJA VISIÓN en pacientes o para determinar su calidad de vida una vez conocido su problema de visión. Gracias al análisis de los DOMINIOS de cada uno de los cuestionarios revisados, se ha obtenido una preselección de 7 de ellos, que van a permitir el desarrollo coherente y consistente de encuestas para estos pacientes.

Capítulo 6

Este capítulo presenta una de las dos principales aportaciones científicas de la presente Tesis Doctoral: el nuevo operador *Convolución-LIP*. Partiendo de la definición matemática del operador *convolución* 2D con filtros separables, y aplicando la teoría del paradigma LIP, se desarrolla una nueva formulación. Al haberse obtenido ésta mediante el mecanismo *Directo* de LIP, se ha dado en llamar DGLIP-CONV. Sin em-

bargo, al implementar esta formulación en procesadores reales, se ha podido observar que su coste computacional es muy elevado. Debido a ello, se ha realizado una nueva formulación que utiliza operadores computacionalmente más sencillos y que proporcionan el mismo resultado pero de manera más rápida. Dicha formulación se ha llamado FGLIP-CONV. Estas dos implementaciones representan una de las aportaciones científicas más relevantes de la Tesis Doctoral, puesto que permiten adaptar a la filosofía LIP cualquier método existente que se base en el uso de la *convolución* con filtros separables, utilizada en la mayoría de técnicas de PROCESAMIENTO DE IMÁGENES de bajo nivel. Se demuestra en dicho capítulo que el rango del nuevo operador es consistente con lo establecido por el paradigma LIP, es decir, el rango de salida se mantiene en $(0, M)$. También se analizan los tiempos de computación, para certificar que en todos los casos el operador FGLIP-CONV es más rápido que DGLIP-CONV. Como nueva aportación de interés, en estas comparativas de tiempos se propone una nueva implementación de FGLIP-CONV, denominada FGLIP-CONV OPTIM., en la que se utiliza el operador clásico *convolución*, que suele estar optimizado para su ejecución en los procesadores actuales. Esto permite obtener resultados aún más rápidamente que con la implementación FGLIP-CONV.

Capítulo 7

Se puede considerar que el Capítulo 7 es en el que se fijan los criterios para el desarrollo concreto del trabajo de investigación. En él se indican los métodos que se seleccionan para su evaluación, las poblaciones muestrales a las que se les pregunta en la encuesta y los DOMINIOS. También, se fijan aspectos de tipo técnico, como el mecanismo de selección de los umbrales y el grosor de los contornos. La principal aportación de este capítulo es el conjunto de imágenes características de cada uno de los DOMINIOS seleccionados, ya que encontrar imágenes que de manera simple representen cada uno de los DOMINIOS principales y que además presenten zonas con sombras (identificado como *Dominio Transversal*), es una tarea ardua.

Capítulo 8

Las recomendaciones para el desarrollo de cuestionarios para la evaluación de la calidad de los métodos de extracción de bordes constituyen otra de las grandes aportaciones científicas de esta Tesis Doctoral. Este capítulo se ha desarrollado como un compendio de propuestas para construir preguntas dentro de cuestionarios con evaluación de contenido visual. Dichas propuestas han sido extraídas del proceso de pilotaje, a partir de los comentarios de los encuestados y de sus respuestas a cuestionarios de prueba. Estos cuestionarios se desarrollaron para analizar las distintas recomendaciones propuestas. También se presenta un sistema Web para el diseño de los cuestionarios con contenidos multimedia, que también sirve para que los encuestados puedan responder a las encuestas diseñadas siguiendo estas recomendaciones. Este sistema Web propuesto también es otra innovación científica.

Capítulo 9

En este capítulo se describen diversos experimentos para mostrar la aplicabilidad del nuevo operador *Convolución*-LIP y su efectividad en los resultados obtenidos. En ellos, se implementan cinco métodos diferentes de dos campos distintos dentro del ámbito del PROCESAMIENTO DE IMÁGENES. Los dos campos de aplicación son la extracción de contornos y el emborronamiento de la imagen. Se exponen los tres algoritmos de extracción de bordes de objetos, tanto en su versión No LIP como en su adaptación a LIP, que se van a evaluar mediante cuestionarios. En general, la adaptación a LIP se realiza mediante tres implementaciones, la tradicional propuesta de **Deng y Pinoli** [29], la nueva propuesta de DGLIP-CONV y la propuesta mejorada FGLIP-CONV. Se demuestra que las diferentes versiones de LIP son equivalentes en cuanto a resultados visuales, aunque la implementación FGLIP-CONV es la más rápida de las tres. En cualquier caso, los resultados visuales de los métodos LIP parecen *a priori* mejores que los obtenidos mediante los algoritmos en su versión No LIP.

Capítulo 10

Este capítulo describe los resultados relativos a las encuestas de evaluación de la calidad de los algoritmos de extracción de contornos. No sólo se indican los resultados estadísticos finales, sino que también se incluyen la descripción del proceso de diseño de los cuestionarios y otros mecanismos que se han utilizado en las diversas fases de creación y utilización de la encuesta. Todo el capítulo en sí se puede considerar una gran novedad investigadora, ya que, hasta donde se conoce en la literatura científica, no se ha descrito con tanto detalle.

12.2 PRODUCCIÓN CIENTÍFICA

La presente Tesis Doctoral ha generado una serie de publicaciones científicas, que se detallan a continuación:

- Rafael Palomar, José M. Palomares, José M. Castillo, Joaquín Olivares y Juan Gómez-Luna *Parallelizing and Optimizing LIP-Canny Using NVIDIA CUDA* En N. García-Pedrajas et al. (editor): *Proceedings of the IEA/AIE 2010, Córdoba (España)* publicado en *Lecture Notes in Artificial Intelligence*, LNAI 6098(III): 389–398, 2010.
- José M. Palomares, Jesús González, Eduardo Ros y Alberto Prieto. *General Logarithmic Image Processing Convolution*. *IEEE Transactions on Image Processing*, 15(11):3602–3608, 2006. DOI: [10.1109/tip.2006.881967](https://doi.org/10.1109/tip.2006.881967). Índice de Impacto: 2.715
- José M. Palomares, Jesús González y Eduardo Ros. *Designing a fast convolution under the LIP paradigm applied to edge detection* En Sameer Singh et al. (editor): *Proceedings of the III ICAPR 2005, Bath (Reino Unido)* publicado en *Lecture Notes on Computer Sciences*, LNCS 3687(3):560–569, 2005. DOI: [10.1007/11552499_62](https://doi.org/10.1007/11552499_62). Índice de Impacto: 0.402
- José M. Palomares, Jesús González y Eduardo Ros. *Detección de bordes en imágenes con sombras mediante LIP-Canny* En las Actas del I Simposio de Reconocimiento de Formas y Análisis de Imágenes: *Conferencia AERFAI 2005, Granada (España)*:71–76, 2005. ISBN:84-9732-445-5

12.3 PRINCIPALES APORTACIONES CIENTÍFICAS

Los objetivos Primarios y Secundarios, inicialmente propuestos en el Capítulo 1, se han satisfecho completamente. A continuación, se indican de manera pormenorizada cada una de las conclusiones obtenidas, que se han agrupado en dos bloques principales para una mejor descripción.

12.3.1 Sobre el operador *Convolución-LIP*

- Hasta el trabajo actual, el paradigma LIP no tenía definido ningún mecanismo de *convolución*. A lo largo de esta Tesis Doctoral, se ha propuesto un nuevo operador *Convolución-LIP*, con dos implementaciones diferentes: DGLIP-CONV y FGLIP-CONV.
- Este nuevo operador se ha desarrollado utilizando una base matemática sólida que ha permitido definirlo de manera generalizada para el uso de cualquier tipo de filtro separable.
- La fundamentación matemática y utilidad de este nuevo operador se ve avalada con la publicación en prestigiosas revistas del ámbito científico mundial [82], en congresos internacionales [79, 80] y en congresos nacionales [81].
- Se ha abstraído y expandido la propuesta original de Deng y Pinoli [29] para poder utilizarla “*ad hoc*” en otros experimentos, a pesar de que los autores originales describieron este método sólo para un experimento en particular. En este sentido, se ha comprobado que dicho método se puede considerar como una particularización optimizada del método DGLIP-CONV.
- Se han adaptado los métodos *LIP-Sobel*, *LIP-LoG* y *LIP-Canny* para la extracción de contornos en imágenes con sombras, mostrando su eficacia visual, su alta velocidad de procesamiento, y su robustez, medida en la cantidad y calidad de contornos detectados y en la facilidad para determinar los umbrales de detección de bordes.
- Se ha desarrollado FGLIP-CONV OPTIM. que es una implementación mejorada del operador *Convolución-LIP*, capaz de utilizar los operadores *convolución* optimizados en los sistemas computacionales actuales. Esta propuesta obtiene mapas de contornos en tiempos de procesamiento muy reducidos (en el orden de los pocos milisegundos).

12.3.2 Sobre las encuestas de evaluación de contornos

- Se han seleccionado un conjunto de imágenes naturales (con cambios de iluminación local) para la evaluación de calidad subjetiva en la extracción de contornos. Este conjunto de imágenes puede servir como base de datos para otros investigadores en la materia.

- Se ha evaluado la calidad de las imágenes procesadas mediante los diferentes métodos propuestos utilizando encuestas de opinión.
- Se ha incluido en este trabajo varios procesos que no son habituales en otros estudios de investigación del ámbito del PROCESAMIENTO DE IMÁGENES, como es el proceso de diseño de dichos cuestionarios, la determinación de la población, los mecanismos de la publicitación de dicha encuesta, el análisis de la muestra poblacional y el estudio concienzudo de los resultados estadísticos obtenidos.
- Se ha propuesto un conjunto de recomendaciones para la construcción de encuestas de evaluación de la calidad de mapas de contornos, basados en la experiencia adquirida en la fase de pre-pilotaje y evaluación previa de los cuestionarios desarrollados.
- Utilizando las recomendaciones descritas con anterioridad se ha desarrollado una encuesta, cuya calidad ha sido refrendada mediante la VALIDACIÓN DE CONTENIDO por 5 expertos (puntuación de 7.8 sobre 10) y mediante el estudio de FIABILIDAD INTERNA, tanto a nivel de *Consistencia interna* (con un α -Cronbach de la encuesta completa = 0.956) como *Grado de acuerdo* (obteniéndose un $ICC(3,k) = 0.993$ para la encuesta completa).
- En las encuestas de opinión se han incluido pacientes con problemas de BAJA VISIÓN, que han evaluado positivamente las imágenes obtenidas mediante el uso del operador *Convulación-LIP*. Estos determinaron que los mapas de contornos generados mediante *LIP-Canny* y *LIP-Sobel* obtenían puntuaciones muy elevadas, lo cual denota que para dicho grupo estos algoritmos tienen buena calidad subjetiva.
- Hasta el momento, no se había determinado si los usuarios con un grado de VISIÓN ESTÁNDAR y los pacientes con BAJA VISIÓN tenían un concepto o constructo mental similar para estimar la calidad de los mapas de contornos. En esta investigación, utilizando los resultados estadísticos extraídos de la encuesta, se ha constatado la similaridad entre la calidad percibida por ambos grupos, pudiéndose afirmar que existe una relación lineal entre la idea de calidad que tienen ambas poblaciones.
- Se ha construido un sistema Web que facilita tanto el diseño de cuestionarios al encuestador como la captación de las respuestas de los encuestados. Este sistema Web permite al encuestador:
 - Acceder a un mayor número de encuestados.
 - Reducir el coste económico.
 - Permitir el seguimiento del trabajo de los encuestados .
 - Facilitar la distribución de los cuestionarios en un entorno mucho más amplio.
- La principal ventaja que presenta el uso del sistema Web para los encuestados es la capacidad de parar y continuar con la encuesta en cualquier instante y en cualquier sitio con conexión a Internet.

12.4 INVESTIGACIONES FUTURAS

El presente trabajo ha proporcionado bases científicas que permite abrir nuevos caminos de investigación. Algunas de las posibles vías de estudio en un futuro se describen a continuación:

- Difundir los resultados de las encuestas de evaluación de contornos en congresos y revistas científicas.
- Implementar el operador *Convolución-LIP* en dispositivos específicos como FPGA o GPU, para obtener rendimientos de tiempo real.
- Diseñar un sistema de ayuda a la visión para personas con BAJA VISIÓN que utilice el algoritmo *LIP-Canny* para la extracción de contornos en entornos reales.
- Construir un mecanismo adaptativo que permita la extracción selectiva de los bordes, de tal forma que se pueda escoger la presentación de los contornos más relevantes en función del modo de navegación escogida por el usuario.
- Utilizando las recomendaciones de diseño propuestas en el Cap. 8, realizar unos cuestionarios más reducidos que permitan aplicar un mecanismo *test-retest*, gracias al cual se pueda comprobar la replicabilidad de los mismos.
- Analizar el uso de imágenes similares en las preguntas de MAPPING, comprobando que los resultados son coherentes con los expuestos en esta Tesis, a la vez que asegura que el *sesgo por finalidad* debido al recuerdo posicional del objeto buscado no se presenta.
- Ampliar el número de pacientes con BAJA VISIÓN, intentando que dicho grupo poblacional sea lo más amplio y variado posible, para comparar los resultados obtenidos en esta Tesis con esos nuevos resultados, intentando garantizar la universalidad de las afirmaciones expuestas.

Además de todas estas posibles vías de investigación futuras, no se descarta el análisis de GANIP [28], que es un nuevo paradigma, altamente emparentado con LIP, que permite utilizar entornos adaptativos en función de las estructuras espaciales utilizadas y de las características intrínsecas de las imágenes procesadas. Este nuevo paradigma está siendo desarrollado en la actualidad desde un punto de vista teórico-matemático por el mismo grupo de autores que desarrollaron el paradigma LIP y podría tener un futuro muy prometedor debido a su capacidad de adaptación espacial del procesamiento.

GANIP: Del inglés, GENERAL ADAPTIVE NEIGHBORHOOD IMAGE PROCESSING, *Procesamiento Generalizado de Entornos Adaptativos de Imágenes.*

Parte V

APÉNDICES

APÉNDICE

ÍNDICE

A.1	Demstración del rango de funcionamiento de la Suma-LIP	239
A.1.1	Valores superiores a cero	240
A.1.2	Valores inferiores a M	240
A.1.3	Rango completo	241

A.1 DEMOSTRACIÓN DEL RANGO DE FUNCIONAMIENTO DE LA SUMA-LIP

A continuación, se demuestra que la aplicación del operador Suma-LIP proporciona resultados en el mismo rango $(0, M)$ que el de los operandos de entrada. Es decir, se desea demostrar:

$$0 < \hat{f} \Delta \hat{g} < M \quad \forall \hat{f}, \hat{g} \in (0, M) \subseteq \mathbb{R} \quad (\text{A.1})$$

Tomando la definición matemática del operador Suma-LIP (definido en (2.12)), se puede reescribir (A.1) utilizando operadores tradicionales.

$$0 < \hat{f} + \hat{g} - \frac{\hat{f} \cdot \hat{g}}{M} < M \quad (\text{A.2})$$

Esta demostración se va a dividir en dos partes. En la primera parte se va a demostrar que $0 < \hat{f} \Delta \hat{g}$; mientras que la segunda parte demostrará que $\hat{f} \Delta \hat{g} < M$.

A.1.1 Valores superiores a cero

Se toma la primera parte que se desea demostrar ($0 < \hat{f} \Delta \hat{g}$). Para confirmar dicha afirmación, se utiliza el mecanismo de *reducción al absurdo*. Para lo cual, se niega la afirmación que se desea demostrar y se opera:

$$0 \not\leq \hat{f} + \hat{g} - \frac{\hat{f} \cdot \hat{g}}{M} \Rightarrow 0 \geq \hat{f} + \hat{g} - \frac{\hat{f} \cdot \hat{g}}{M} \quad (\text{A.3})$$

Para simplificar las operaciones, se va a utilizar el mismo operando. Esta simplificación no significa, en ningún caso, una falta de generalidad en esta demostración, ya que puede ampliarse a dos operandos cualesquiera. Se reformula (A.3) utilizando un mismo operando, \hat{f} , y se continúa operando:

$$\begin{aligned} 0 &\geq \hat{f} + \hat{f} - \frac{\hat{f} \cdot \hat{f}}{M} \\ 0 &\geq 2 \cdot \hat{f} - \frac{\hat{f}^2}{M} \Rightarrow 0 \geq 2 \cdot M \cdot \hat{f} - \hat{f}^2 \\ \frac{\hat{f}^2}{M} &\geq 2 \cdot \hat{f} \Rightarrow \frac{\hat{f}}{2} \geq M \end{aligned} \quad (\text{A.4})$$

Sabiendo que por definición, $\hat{f} \in (0, M) \subseteq \mathbb{R}$, a continuación, se calculan los límites de los valores extremos del rango de \hat{f} para (A.4).

$$\lim_{\hat{f} \rightarrow 0} \left(\frac{\hat{f}}{2} \right) = 0 \quad (\text{A.5})$$

$$\lim_{\hat{f} \rightarrow M} \left(\frac{\hat{f}}{2} \right) = \frac{M}{2} \quad (\text{A.6})$$

Para (A.5), el análisis es trivial, ya que el único caso para el que $0 \geq M$, se produce cuando $M = 0$. Esto representaría un rango de trabajo nulo, aspecto que no tiene sentido tanto desde el punto de vista matemático como desde el punto de vista práctico.

La demostración para (A.6) también es trivial, puesto que $\frac{M}{2} \geq M$ es válido únicamente si $M = 0$, que como se ha indicado anteriormente, no es matemáticamente posible.

Con lo cual, se puede afirmar con total seguridad que la hipótesis de partida es falsa, y por tanto:

$$0 < \hat{f} \Delta \hat{g} \quad (\text{A.7})$$

A.1.2 Valores inferiores a M

La segunda parte de (A.2) pretende demostrar que ($\hat{f} \Delta \hat{g} < M$). Para dicha demostración, como en el caso anterior, se utiliza el mecanismo

de *reducción al absurdo*. Se niega la afirmación que se desea demostrar y se opera:

$$\widehat{f} + \widehat{g} - \frac{\widehat{f} \cdot \widehat{g}}{M} < M \Rightarrow \widehat{f} + \widehat{g} - \frac{\widehat{f} \cdot \widehat{g}}{M} \geq M \quad (\text{A.8})$$

Como en el caso anterior, para simplificar las operaciones, se realizan los cálculos utilizando el mismo operando. Por lo que, reescribiendo (A.8) con un mismo operando, \widehat{f} , se obtiene:

$$\begin{aligned} \widehat{f} + \widehat{f} - \frac{\widehat{f} \cdot \widehat{f}}{M} &\geq M \\ 2 \cdot \widehat{f} - \frac{\widehat{f}^2}{M} &\geq M \Rightarrow 2 \cdot M \cdot \widehat{f} - \widehat{f}^2 \geq M^2 \\ 0 &\geq M^2 - 2 \cdot M \cdot \widehat{f} + \widehat{f}^2 \Rightarrow 0 \geq (M - \widehat{f})^2 \\ 0 &\geq M - \widehat{f} \Rightarrow \widehat{f} \geq M \end{aligned} \quad (\text{A.9})$$

Sin embargo, por definición, $\widehat{f} \in (0, M) \subseteq \mathbb{R}$. Por tanto, se tiene que $\widehat{f} < M$, lo que hace que la inecuación (A.9) no sea válida en ningún caso. Esto permite deducir que el punto de partida es falso, y se puede afirmar lo contrario:

$$\widehat{f} \triangle \widehat{g} < M \quad (\text{A.10})$$

A.1.3 Rango completo

Al unificar (A.7) y (A.10) queda demostrado que el rango de salida del operador Suma-LIP es $(0, M)$ para cualesquiera dos operandos $\widehat{f}, \widehat{g} \in (0, M) \subseteq \mathbb{R}$, tal y como se ha expuesto en (A.1).

ÍNDICE

B.1	Demostración del rango de funcionamiento de la Multiplicación-LIP	243
B.1.1	Valores superiores a cero	244
B.1.2	Valores inferiores a M	244
B.1.3	Rango completo	245

B.1 DEMOSTRACIÓN DEL RANGO DE FUNCIONAMIENTO DE LA MULTIPLICACIÓN-LIP

En este apéndice se demuestra que el rango de los resultados del operador Multiplicación-LIP es $(0, M)$. Es decir, se desea demostrar:

$$0 < \alpha \Delta \hat{f} < M \quad \forall \alpha \in \mathbb{R}, \forall \hat{f} \in (0, M) \subseteq \mathbb{R} \quad (\text{B.1})$$

Se reescribe (B.1) a partir de la definición matemática del operador Multiplicación-LIP (definido en (2.13)), utilizando únicamente operadores tradicionales.

$$0 < M - M \cdot \left(1 - \frac{\hat{f}}{M}\right)^\alpha < M \quad (\text{B.2})$$

Esta demostración se divide en dos trozos. En el primero, se demuestra que $0 < \alpha \Delta \hat{f}$; mientras que en una segunda parte se demuestra que $\alpha \Delta \hat{f} < M$.

B.1.1 Valores superiores a cero

Se toma la primera parte que se desea demostrar ($0 < \alpha \Delta \hat{f}$). Para confirmar dicha afirmación, se utiliza el mecanismo de *reducción al absurdo*. Para lo cual, se niega la afirmación que se desea demostrar:

$$0 \not\leq M - M \cdot \left(1 - \frac{\hat{f}}{M}\right)^\alpha \Rightarrow 0 \geq M - M \cdot \left(1 - \frac{\hat{f}}{M}\right)^\alpha \quad (\text{B.3})$$

Partiendo de (B.3), se opera hasta obtener inecuación simple que permita un análisis directo.

$$\begin{aligned} 0 \geq M \cdot \left(1 - \left(\frac{M - \hat{f}}{M}\right)^\alpha\right) &\Rightarrow 0 \geq 1 - \left(\frac{M - \hat{f}}{M}\right)^\alpha \\ \left(\frac{M - \hat{f}}{M}\right)^\alpha &\geq 1 \Rightarrow \frac{(M - \hat{f})^\alpha}{M^\alpha} \geq 1 \\ (M - \hat{f})^\alpha &\geq M^\alpha \end{aligned} \quad (\text{B.4})$$

Se aplica el logaritmo neperiano a ambas partes de la inecuación (B.4).

$$\begin{aligned} \log(M - \hat{f})^\alpha &\geq \log(M^\alpha) \Rightarrow \alpha \cdot \log(M - \hat{f}) \geq \alpha \cdot \log M \\ \log(M - \hat{f}) &\geq \log M \Rightarrow \frac{\log(M - \hat{f})}{\log M} \geq 0 \\ \log(M - \hat{f} - M) &\geq 0 \Rightarrow \log(-\hat{f}) \geq 0 \end{aligned} \quad (\text{B.5})$$

Sabiendo que, por definición, $\hat{f} \in (0, M) \subseteq \mathbb{R}$, la inecuación (B.5) no puede satisfacerse, puesto que no existe el logaritmo de un número negativo. Consecuentemente, se puede afirmar con total seguridad que la hipótesis de partida es falsa, y por tanto:

$$0 < \alpha \Delta \hat{f} \quad \forall \alpha \in \mathbb{R}, \forall \hat{f} \in (0, M) \subseteq \mathbb{R} \quad (\text{B.6})$$

Nótese que no influye el valor de α en la determinación del rango de salida.

B.1.2 Valores inferiores a M

La segunda parte de (B.2) pretende demostrar que ($\alpha \Delta \hat{f} < M$). Para dicha demostración, como en el caso anterior, se utiliza el mecanismo de *reducción al absurdo*. Se niega la afirmación que se desea demostrar y se opera:

$$M - M \cdot \left(1 - \frac{\hat{f}}{M}\right)^\alpha \not< M \Rightarrow M - M \cdot \left(1 - \frac{\hat{f}}{M}\right)^\alpha \geq M \quad (\text{B.7})$$

Como en el caso anterior, para simplificar las operaciones, se realizan los cálculos utilizando el mismo operando. Por lo que, reescribiendo (A.8) con un mismo operando, \hat{f} , se obtiene:

$$\begin{aligned} M \cdot \left(1 - \left(1 - \frac{\hat{f}}{M}\right)^\alpha\right) \geq M &\Rightarrow 1 - \left(1 - \frac{\hat{f}}{M}\right)^\alpha \geq 1 \\ - \left(1 - \frac{\hat{f}}{M}\right)^\alpha \geq 0 &\Rightarrow \left(1 - \frac{\hat{f}}{M}\right)^\alpha \leq 0 \\ (M - \hat{f})^\alpha \leq 0 &\Rightarrow M \cdot \hat{f} \leq 0 \Rightarrow M \leq \hat{f} \quad (\text{B.8}) \end{aligned}$$

Sin embargo, por definición, $\hat{f} \in (0, M) \subseteq \mathbb{R}$. Por tanto, se tiene que $\hat{f} < M$, lo que hace que la inecuación (A.9) no sea cierta. Esto permite deducir que el punto de partida es falso, y se puede afirmar lo contrario:

$$\alpha \Delta \hat{f} < M \quad \forall \alpha \in \mathbb{R}, \forall \hat{f} \in (0, M) \subseteq \mathbb{R} \quad (\text{B.9})$$

B.1.3 Rango completo

Al unificar (B.6) y (B.9) queda demostrado que el rango de salida del operador Multiplicación-LIP es $(0, M)$ para cualquier valor del operando $\alpha \in \mathbb{R}$, $\hat{f} \in (0, M) \subseteq \mathbb{R}$, tal y como se ha expuesto en (B.1).

APÉNDICE

ÍNDICE

c.1	Cuestionarios utilizados en el estudio	248
c.1.1	Cuestionario 1	248
c.1.2	Cuestionario 2	254

El estudio de opinión de calidad percibida de los diferentes algoritmos de extracción de contornos se ha realizado mediante dos cuestionarios. En este apéndice, se incluyen los enunciados de cada una de las 22 preguntas de cada uno de los dos cuestionarios. También se incluyen las *imágenes respuesta* incluidas en cada pregunta. Se ha decidido no incluir gráficamente dichas *imágenes respuesta* en la memoria de esta Tesis Doctoral, debido al número tan elevado (244 *imágenes respuesta*) utilizado. La inclusión de las mismas significaría un aumento excesivo del tamaño de esta memoria de investigación, presentado por el contrario una utilidad muy reducida. Todas estas *imágenes respuesta* están disponibles en <http://www.uco.es/users/el2pamuj/research/LowVision/>.

C.1 CUESTIONARIOS UTILIZADOS EN EL ESTUDIO

c.1.1 Cuestionario 1

1. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar una motocicleta en sombras utilizando cada imagen de contornos, donde 0 significa que no es posible detectar la motocicleta en absoluto con dicha imagen de contornos y 10 representa que es posible detectar dicha motocicleta completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Exteriores 2 - calle_sombra5
 - Exteriores 2 - Canny 80
 - Exteriores 2 - Sobel 80
 - Exteriores 2 - LIPCanny 80
 - Exteriores 2 - LoG 95
 - Exteriores 2 - LIPSobel 80
 - Exteriores 2 - LIPLoG 95

2. En las siguientes imágenes de contornos aparecen personas. Puntúe cada imagen de bordes de 0 a 10, en función de la calidad de las mismas para reconocer los rasgos de las caras y los posibles gestos que estén realizando. Una puntuación de 0 significa que no es posible reconocer ningún rasgo de la cara ni es posible apreciar ningún rasgo; la puntuación de 10 significa que se reconocen todos los rasgos de la cara, pudiendo apreciar completamente el gesto de la cara.
 - Personas 1 - Canny 80
 - Personas 1 - LIPSobel 80
 - Personas 1 - LIP-Canny 80
 - Personas 1 - LoG 95
 - Personas 1 - Sobel 80
 - Personas 1 - LIPLoG 95

3. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
 - Exteriores 2 - Sobel 80

4. En la siguiente imagen se muestra una escena habitual de la vida diaria, en la que se utilizan diferentes objetos comunes. Puntúe de 0 a 10, cada una de las imágenes de bordes en función de la calidad de los contornos para detectar todos los objetos presentes en la imagen original. Una puntuación de 0 representa que no

es posible detectar ningún objeto, mientras que una puntuación de 10 indica que todos los objetos en la imagen han sido detectados y que los contornos de todos los objetos han sido detectados completamente.

- Cercanas 1 - alacena
 - Cercanas 1 - Sobel 80
 - Cercanas 1 - LoG 95
 - Cercanas 1 - LIPCanny 80
 - Cercanas 1 - LIPLoG 95
 - Cercanas 1 - Canny 80
 - Cercanas 1 - LIPSobel 80
5. A continuación se muestra una imagen de una casa. Dicha casa tiene 6 ventanas (2 en la zona iluminada y 4 en la zona en sombras). Valore de 0 a 10 cada imagen de contornos, sin repetir la puntuación, la calidad de cada imagen de contornos para identificar claramente las 6 ventanas. La puntuación de 0 significa que la imagen de contornos no permite identificar las todas las ventanas, mientras que 10 permite identificar todas las ventanas completamente, sin que les falte detalle y no existen trazos que molesten la identificación.
- Exteriores 3 - casa
 - Exteriores 3 - Sobel 80
 - Exteriores 3 - LIPSobel 80
 - Exteriores 3 - LIPLoG 95
 - Exteriores 3 - LIPCanny 80
 - Exteriores 3 - LoG 95
 - Exteriores 3 - Canny 80
6. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
- Exteriores 2 - LIPCanny 80
7. En las siguientes imágenes de contorno hay un grupo de personas en una terraza de un bar que se encuentra en una zona de sombras. Indique en cuál de las imágenes puede distinguir a las distintas personas en el bar. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Exteriores 1 - LIPCanny 80
 - Exteriores 1 - LIPLoG 95
 - Exteriores 1 - Canny 80
 - Exteriores 1 - LIPSobel 80
 - Exteriores 1 - Sobel 80

- Exteriores 1 - LoG 95
8. En las siguientes imágenes de contornos aparecen personas. Puntúe cada imagen de bordes de 0 a 10, en función de la calidad de las mismas para reconocer los rasgos de las caras y los posibles gestos que estén realizando. Una puntuación de 0 significa que no es posible reconocer ningún rasgo de la cara ni es posible apreciar ningún rasgo; la puntuación de 10 significa que se reconocen todos los rasgos de la cara, pudiendo apreciar completamente el gesto de la cara.
 - Personas 1 - Canny 80
 - Personas 1 - Sobel 80
 - Personas 1 - LoG 95
 - Personas 1 - LIPLoG 95
 - Personas 1 - LIPSobel 80
 - Personas 1 - LIP-Canny 80
 9. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
 - Exteriores 2 - LoG 95
 10. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar las ventanas bajo los soportales utilizando cada imagen de contornos, donde 0 significa que no es posible detectar las ventanas bajo los soportales en absoluto con dicha imagen de contornos y 10 representa que es posible detectar las ventanas bajo los soportales completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Exteriores 4 - soportales
 - Exteriores 4 - Sobel 80
 - Exteriores 4 - Canny 80
 - Exteriores 4 - LIPLoG 95
 - Exteriores 4 - LIPSobel 80
 - Exteriores 4 - LoG 95
 - Exteriores 4 - LIPCanny 80
 11. A continuación se muestra la imagen real de la que se han obtenido las siguientes imágenes de contornos. Indique en cuáles de dichas imágenes de contornos puede Vd. detectar los gestos del vendedor que se encuentra agachado y en sombras. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Personas 5 - vendedor-sombras

- Personas 5 - LIPSobel 80
 - Personas 5 - Sobel 80
 - Personas 5 - LIPCanny 80
 - Personas 5 - LoG 95
 - Personas 5 - Canny 80
 - Personas 5 - LIPLoG 95
12. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
- Exteriores 2 - Canny 80
13. En las siguientes imágenes de contornos aparecen personas. Puntúe cada imagen de bordes de 0 a 10, en función de la calidad de las mismas para reconocer los rasgos de las caras y los posibles gestos que estén realizando. Una puntuación de 0 significa que no es posible reconocer ningún rasgo de la cara ni es posible apreciar ningún rasgo; la puntuación de 10 significa que se reconocen todos los rasgos de la cara, pudiendo apreciar completamente el gesto de la cara.
- Personas 4 - LIPLoG 95
 - Personas 4 - LIPSobel 80
 - Personas 4 - Canny 80
 - Personas 4 - Sobel 80
 - Personas 4 - LIPCanny 80
 - Personas 4 - LoG 95
14. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar todos los objetos presentes en las estanterías utilizando cada imagen de contornos, donde 0 significa que no es posible detectar completamente ningún objeto en absoluto con dicha imagen de contornos y 10 representa que es posible detectar absolutamente todos los objetos completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Cercanas 2 - despensa
 - Cercanas 2 - Sobel 80
 - Cercanas 2 - LoG 95
 - Cercanas 2 - LIPCanny 80
 - Cercanas 2 - LIPSobel 80
 - Cercanas 2 - LIPLoG 95
 - Cercanas 2 - Canny 80

15. En las siguientes imágenes de contorno hay un coche con una de las ruedas traseras que se encuentra en una zona de sombras. Indique en cuál de las imágenes puede distinguir perfecta y claramente dicha rueda trasera. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Exteriores 5 - Sobel 80
 - Exteriores 5 - LIPCanny 80
 - Exteriores 5 - LoG 95
 - Exteriores 5 - Canny 80
 - Exteriores 5 - LIPSobel 80
 - Exteriores 5 - LIPLoG 95
16. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto.
 - Exteriores 2 - LIPSobel 80
17. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar un camino libre de obstáculos para atravesar el pasillo sin chocar utilizando cada imagen de contornos, donde 0 significa que no es posible detectar encontrar un camino que permitiese recorrer con seguridad dicho pasillo en absoluto con dicha imagen de contornos y 10 representa que es posible detectar ese recorrido seguro evitando los obstáculos completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Interiores 1 - pasillo-cuadros2
 - Interiores 1 - LIPSobel 80
 - Interiores 1 - LIPCanny 80
 - Interiores 1 - LIPLoG 95
 - Interiores 1 - Canny 90
 - Interiores 1 - LoG 95
 - Interiores 1 - Sobel 80
18. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar todos los ventanales de esta galería utilizando cada imagen de contornos, donde 0 significa que no es posible detectar los ventanales en absoluto con dicha imagen de contornos y 10 representa que es posible detectar los ventanales completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Interiores 2 - galeria-budapest
 - Interiores 2 - LoG 95

- Interiores 2 - Sobel 80
 - Interiores 2 - LIPCanny 80
 - Interiores 2 - LIPLoG 95
 - Interiores 2 - LIPSobel 80
 - Interiores 2 - Canny 80
19. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
- Exteriores 2 - LIPLoG 95
20. A continuación se muestra la imagen real de la que se han obtenido las siguientes imágenes de contornos. Indique en cuáles de dichas imágenes de contornos puede Vd. detectar la puerta de salida al final del pasillo. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Interiores 3 - pasillo1-noluz
 - Interiores 3 - LIPLoG 95
 - Interiores 3 - LIPCanny 80
 - Interiores 3 - Sobel 80
 - Interiores 3 - LoG 95
 - Interiores 3 - Canny 80
 - Interiores 3 - LIPSobel 80
21. A continuación se muestra la imagen real de la que se han obtenido las siguientes imágenes de contornos. Indique en cuáles de dichas imágenes de contornos puede Vd. detectar mejor los rasgos de la cara. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Personas 2 - cara-gorra
 - Personas 2 - LIPLoG 95
 - Personas 2 - LIPSobel 80
 - Personas 2 - Sobel 80
 - Personas 2 - LIP-Canny 80
 - Personas 2 - Canny 80
 - Personas 2 - LoG 95
22. En las siguientes imágenes de contorno hay un mendigo en cuclillas que se encuentra en una zona de sombras. Indique en cuál de las imágenes puede distinguir mejor todos los rasgos de la cara de dicha persona. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Personas 3 - LoG 95
 - Personas 3 - LIP-Canny 80

- Personas 3 - Sobel 80
- Personas 3 - LIPLoG 95
- Personas 3 - Canny 80
- Personas 3 - LIPSobel 80

c.1.2 Cuestionario 2

1. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar una motocicleta en sombras utilizando cada imagen de contornos, donde 0 significa que no es posible detectar la motocicleta en absoluto con dicha imagen de contornos y 10 representa que es posible detectar dicha motocicleta completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Exteriores 2 - calle_sombra5
 - Exteriores 2 - LIPSobel 90
 - Exteriores 2 - LIPLoG 120
 - Exteriores 2 - LoG 120
 - Exteriores 2 - Sobel 90
 - Exteriores 2 - Canny 90
 - Exteriores 2 - LIPCanny 90
2. En las siguientes imágenes de contornos aparecen personas. Puntúe cada imagen de bordes de 0 a 10, en función de la calidad de las mismas para reconocer los rasgos de las caras y los posibles gestos que estén realizando. Una puntuación de 0 significa que no es posible reconocer ningún rasgo de la cara ni es posible apreciar ningún rasgo; la puntuación de 10 significa que se reconocen todos los rasgos de la cara, pudiendo apreciar completamente el gesto de la cara.
 - Personas 1 - LoG 120
 - Personas 1 - LIP-Canny 90
 - Personas 1 - LIPSobel 90
 - Personas 1 - Canny 90
 - Personas 1 - LIPLoG 120
 - Personas 1 - Sobel 90
3. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
 - Exteriores 2 - Sobel 90

4. En la siguiente imagen se muestra una escena habitual de la vida diaria, en la que se utilizan diferentes objetos comunes. Puntúe de 0 a 10, cada una de las imágenes de bordes en función de la calidad de los contornos para detectar todos los objetos presentes en la imagen original. Una puntuación de 0 representa que no es posible detectar ningún objeto, mientras que una puntuación de 10 indica que todos los objetos en la imagen han sido detectados y que los contornos de todos los objetos han sido detectados completamente.
 - Cercanas 1 - alacena
 - Cercanas 1 - LoG 120
 - Cercanas 1 - Sobel 90
 - Cercanas 1 - LIPCanny 90
 - Cercanas 1 - LIPSobel 90
 - Cercanas 1 - Canny 90
 - Cercanas 1 - LIPLoG 120

5. A continuación se muestra una imagen de una casa. Dicha casa tiene 6 ventanas (2 en la zona iluminada y 4 en la zona en sombras). Valore de 0 a 10 cada imagen de contornos, sin repetir la puntuación, la calidad de cada imagen de contornos para identificar claramente las 6 ventanas. La puntuación de 0 significa que la imagen de contornos no permite identificar las todas las ventanas, mientras que 10 permite identificar todas las ventanas completamente, sin que les falte detalle y no existen trazos que molesten la identificación.
 - Exteriores 3 - casa
 - Exteriores 3 - Sobel 90
 - Exteriores 3 - LIPLoG 120
 - Exteriores 3 - Canny 90
 - Exteriores 3 - LoG 120
 - Exteriores 3 - LIPCanny 90
 - Exteriores 3 - LIPSobel 90

6. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
 - Exteriores 2 - LIPLoG 120

7. En las siguientes imágenes de contorno hay un grupo de personas en una terraza de un bar que se encuentra en una zona de sombras. Indique en cuál de las imágenes puede distinguir a las distintas personas en el bar. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Exteriores 1 - LIPCanny 90

- Exteriores 1 - Sobel 90
 - Exteriores 1 - LoG 120
 - Exteriores 1 - LIPLoG 120
 - Exteriores 1 - Canny 90
 - Exteriores 1 - LIPSobel 90
8. En las siguientes imágenes de contornos aparecen personas. Puntúe cada imagen de bordes de 0 a 10, en función de la calidad de las mismas para reconocer los rasgos de las caras y los posibles gestos que estén realizando. Una puntuación de 0 significa que no es posible reconocer ningún rasgo de la cara ni es posible apreciar ningún rasgo; la puntuación de 10 significa que se reconocen todos los rasgos de la cara, pudiendo apreciar completamente el gesto de la cara.
- Personas 1 - Sobel 90
 - Personas 1 - LoG 120
 - Personas 1 - LIPSobel 90
 - Personas 1 - LIP-Canny 90
 - Personas 1 - Canny 90
 - Personas 1 - LIPLoG 120
9. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
- Exteriores 2 - Canny 90
10. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar las ventanas bajo los soportales utilizando cada imagen de contornos, donde 0 significa que no es posible detectar las ventanas bajo los soportales en absoluto con dicha imagen de contornos y 10 representa que es posible detectar las ventanas bajo los soportales completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Exteriores 4 - soportales
 - Exteriores 4 - Canny 90
 - Exteriores 4 - LIPSobel 90
 - Exteriores 4 - LoG 120
 - Exteriores 4 - LIPLoG 120
 - Exteriores 4 - LIPCanny 90
 - Exteriores 4 - Sobel 90

11. A continuación se muestra la imagen real de la que se han obtenido las siguientes imágenes de contornos. Indíque en cuáles de dichas imágenes de contornos puede Vd. detectar los gestos del vendedor que se encuentra agachado y en sombras. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Personas 5 - vendedor-sombras
 - Personas 5 - LIPCanny 90
 - Personas 5 - Sobel 90
 - Personas 5 - Canny 90
 - Personas 5 - LIPLoG 120
 - Personas 5 - LIPSobel 90
 - Personas 5 - LoG 120

12. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.
 - Exteriores 2 - LIPSobel 90

13. En las siguientes imágenes de contornos aparecen personas. Puntúe cada imagen de bordes de 0 a 10, en función de la calidad de las mismas para reconocer los rasgos de las caras y los posibles gestos que estén realizando. Una puntuación de 0 significa que no es posible reconocer ningún rasgo de la cara ni es posible apreciar ningún rasgo; la puntuación de 10 significa que se reconocen todos los rasgos de la cara, pudiendo apreciar completamente el gesto de la cara.
 - Personas 4 - LoG 120
 - Personas 4 - Canny 90
 - Personas 4 - LIPLoG 120
 - Personas 4 - Sobel 90
 - Personas 4 - LIPSobel 90
 - Personas 4 - LIPCanny 90

14. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar todos los objetos presentes en las estanterías utilizando cada imagen de contornos, donde 0 significa que no es posible detectar completamente ningún objeto en absoluto con dicha imagen de contornos y 10 representa que es posible detectar absolutamente todos los objetos completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
 - Cercanas 2 - despensa
 - Cercanas 2 - LIPLoG 120

- Cercanas 2 - LIPCanny 90
 - Cercanas 2 - Sobel 90
 - Cercanas 2 - LIPSobel 90
 - Cercanas 2 - Canny 90
 - Cercanas 2 - LoG 120
15. En las siguientes imágenes de contorno hay un coche con una de las ruedas traseras que se encuentra en una zona de sombras. Indique en cuál de las imágenes puede distinguir perfecta y claramente dicha rueda trasera. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Exteriores 5 - LIPSobel 90
 - Exteriores 5 - LIPCanny 90
 - Exteriores 5 - Canny 90
 - Exteriores 5 - LoG 120
 - Exteriores 5 - LIPLoG 120
 - Exteriores 5 - Sobel 90
16. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto.
- Exteriores 2 - LIPCanny 90
17. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar un camino libre de obstáculos para atravesar el pasillo sin chocar utilizando cada imagen de contornos, donde 0 significa que no es posible detectar encontrar un camino que permitiese recorrer con seguridad dicho pasillo en absoluto con dicha imagen de contornos y 10 representa que es posible detectar ese recorrido seguro evitando los obstáculos completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Interiores 1 - pasillo-cuadros2
 - Interiores 1 - Canny 90
 - Interiores 1 - LoG 120
 - Interiores 1 - LIPCanny 90
 - Interiores 1 - LIPSobel 90
 - Interiores 1 - Sobel 90
 - Interiores 1 - LIPLoG 120
18. En las siguientes imágenes de contornos se muestran los bordes de la imagen de referencia, mostrada a continuación. Puntúe las imágenes de contornos de 0 a 10 en función de la capacidad de detectar todos los ventanales de esta galería utilizando cada imagen de contornos, donde 0 significa que no es posible detectar

los ventanales en absoluto con dicha imagen de contornos y 10 representa que es posible detectar los ventanales completamente, pudiéndose reconocer todas las características de dicho objeto. (Recuerde pinchar en las imágenes para verlas en su tamaño original)

- Interiores 2 - galeria-budapest
- Interiores 2 - LIPSobel 90
- Interiores 2 - LIPLoG 120
- Interiores 2 - Canny 90
- Interiores 2 - Sobel 90
- Interiores 2 - LoG 120
- Interiores 2 - LIPCanny 90

19. A continuación se mostrará una imagen de contornos en la que hay un coche aparcado en sombras. Pinche con el ratón donde Vd. considere que se encuentre el centro del objeto. Si Vd. no encuentra dicho objeto en la imagen, pulse en la esquina superior izquierda en la imagen.

- Exteriores 2 - LoG 120

20. A continuación se muestra la imagen real de la que se han obtenido las siguientes imágenes de contornos. Indíque en cuáles de dichas imágenes de contornos puede Vd. detectar la puerta de salida al final del pasillo. (Recuerde pinchar en las imágenes para verlas en su tamaño original)

- Interiores 3 - pasillo1-noluz
- Interiores 3 - Canny 90
- Interiores 3 - Sobel 90
- Interiores 3 - LIPCanny 90
- Interiores 3 - LIPLoG 120
- Interiores 3 - LoG 120
- Interiores 3 - LIPSobel 90

21. A continuación se muestra la imagen real de la que se han obtenido las siguientes imágenes de contornos. Indique en cuáles de dichas imágenes de contornos puede Vd. detectar mejor los rasgos de la cara. (Recuerde pinchar en las imágenes para verlas en su tamaño original)

- Personas 2 - cara-gorra
- Personas 2 - LIPLoG 120
- Personas 2 - Sobel 90
- Personas 2 - LIP-Canny 90
- Personas 2 - LIPSobel 90
- Personas 2 - LoG 120
- Personas 2 - Canny 90

22. En las siguientes imágenes de contorno hay un mendigo en cuclillas que se encuentra en una zona de sombras. Indique en cuál de las imágenes puede distinguir mejor todos los rasgos de la cara de dicha persona. (Recuerde pinchar en las imágenes para verlas en su tamaño original)
- Personas 3 - Sobel 90
 - Personas 3 - LIP-Canny 90
 - Personas 3 - LoG 120
 - Personas 3 - LIPSobel 90
 - Personas 3 - LIPLoG 120
 - Personas 3 - Canny 90

BIBLIOGRAFÍA

- [1] Ikram E. Abdou and Nicolas J. Dusaussoy: *Survey of image measurements*. In *Proceedings of the 1986 ACM Fall Joint Computer Conference*, pages 71–78, Dallas, TX, USA, 1986. IEEE Computer Society Press. http://portal.acm.org/ft_gateway.cfm?id=324546&type=pdf&coll=&dl=ACM&CFID=15151515&CFTOKEN=6184618. (Citado en la página 120.)
- [2] F. Alvira-Martín: *La Encuesta: Una perspectiva general metodológica*. Ed. Centro de Investigaciones Sociológicas, 2004. (Citado en las páginas 70, 72 y 73.)
- [3] H. Apfelbaum, D. Apfelbaum, R. Woods, and Eli Peli: *The effect of edge filtering on vision multiplexing*. *SID Digest*, 1:5, 2005. (Citado en la página 10.)
- [4] J.M. Argimon-Pallás y J. Jiménez-Villa: *Métodos de investigación clínica y epidemiológica*. Elsevier, 3^a edición, 2005. (Citado en la página 73.)
- [5] Baja Visión Angel Barañano: *Dossier sobre la Baja Visión*. Informe técnico, BVAB., 2007. <http://www.baja-vision.org/dossier.pdf>. (Citado en las páginas 4 y 8.)
- [6] Baja Visión Angel Barañano: *¿Qué es la Baja Visión?* Website, Agosto 2007. <http://www.baja-vision.org/bajavision.htm>. (Citado en las páginas 8 y 9.)
- [7] D. A. Baylor, B.J. Nunn, and J.L. Schnapf: *The photocurrent, noise and spectral sensitivity of rods of the monkey macaca fascicularis*. *Journal of Physiology*, 347(1):575–607, December 1984. (Citado en la página 30.)
- [8] B Bergman and J Sjöstrand: *Vision and visual disability in the daily life of a representative population sample aged 82 years*. *Acta Ophthalmologica*, 70:33–43, 1992. (Citado en la página 93.)
- [9] P. Bernth-Petersen: *Visual functioning in cataract patients: Methods of measuring and results*. *Acta Ophthalmologica*, 59:198–205, 1981. (Citado en la página 90.)
- [10] TH Bikson and TK Bikson: *Functional problems of the visually impaired: a research approach*. Rand paper series, p-6648, Santa Monica, CA, Rand Corporation, 1981. (Citado en la página 91.)
- [11] Jay David Bolter and Blair Macintyre: *Is it live or is it AR?* *IEEE Spectrum (INT)*, 44(8):24–29, August 2007. (Citado en la página 12.)

- [12] Al Bovik (editor): *Handbook of Image and Video Processing*. ISBN: 0-12-119790-5. Academic Press, San Diego (CA), USA and London, UK., 1st edition, 2000. (Citado en las páginas 33, 34, 51 y 120.)
- [13] Norman M. Bradburn, Brian Wansink, and Seymour Sudman: *Asking questions: the definitive guide to questionnaire design — for market research, political polls, and social and health questionnaires*. ISBN: 0-7879-7088-3. John Wiley & Sons, Inc., San Francisco, USA., 1st edition, 2004. (Citado en la página 76.)
- [14] MH Brenner, B Curbow, and JC Javitt: *Vision change and quality of life in the elderly. response to cataract surgery and treatment of other chronic ocular conditions*. Arch Ophthalmol, 111:680-685, 1993. (Citado en la página 94.)
- [15] B. Bringier, L. Quintard, and M. C. Larabi: *Quality assesement for CRT and LCD color reproduction using a blind metric*. Electronic Letters on Computer Vision and Image Analysis, 7(3):23-25, 2008. (Citado en la página 54.)
- [16] Thomas F. Burgess: *A general introduction to the design of questionnaires for survey research*. Technical report, University of Leeds, May 2001. (Citado en las páginas 64 y 118.)
- [17] John Canny: *A computational approach to edge detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679-698, November 1986. (Citado en las páginas 6, 35, 41 y 169.)
- [18] M. Carnec, P. Le Callet, and D. Barba: *Full reference and reduced reference metrics for image quality assessment*. In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, volume 1, pages 477-480 vol.1, July 2003. (Citado en la página 49.)
- [19] CCIR: *Test pictures and sequences for subjective assessments of digital codecs*. Report 1213, 1990. Anex to Volume IX, Part 1. (Citado en la página 56.)
- [20] M.A. Cea D'ancona: *Métodos de encuesta. Teoría y práctica, errores y mejora*. Ed. Síntesis, Madrid, 2004. (Citado en las páginas 71, 72 y 73.)
- [21] Clarity: *Products*. Website, March 2009. http://www.clarityusa.com/index.php?option=com_content&task=view&id=17&Itemid=35. (Citado en la página 4.)
- [22] Walter Cohen: *Color-perception in the chromatic Ganzfeld*. The American Journal of Psychology, 71(2):390-394, June 1958. (Citado en la página 51.)
- [23] Consejo Argentino de Oftalmología: *Cuestionario de Calidad Visual (CCV-07) auto evaluativo-asistido*. Website, Junio 2008. http://www.bajavision.org.ar/pacientes/evalua_calidad.html. (Citado en la página 97.)

- [24] Coren, Ward, and Enns: *Sensation and Perception*. Hartcourt Brace, 5th edition, 1998. (Citado en la página 51.)
- [25] S Coren and R Hakstian: *Visual screening without the use of technical equipment: preliminary development of a behaviorally validated questionnaire*. *Journal of Applied Optics*, 26:1468–1472, 1987. (Citado en la página 92.)
- [26] COST 211quat: *Call for Analysis Model Comparisons*. Website, 2001. <http://www.iva.cs.tut.fi/COST211/Call/Call.htm>. (Citado en la página 56.)
- [27] DARPA: *Grand challenge*. Website. <http://www.darpa.mil/grandchallenge/>. (Citado en la página 4.)
- [28] Johan Debayle, , and Jean Charles Pinoli: *General adaptive neighborhood choquet image filtering*. *Journal of Mathematical Imaging and Vision*, 35(3):173–185, 2009. (Citado en la página 235.)
- [29] Guang Deng and Jean Charles Pinoli: *Differentiation-based edge detection using the logarithmic image processing model*. *Journal of Mathematical Imaging and Vision*, 8(2):161–180, March 1998. (Citado en las páginas 17, 28, 158, 159, 160, 180, 220, 221, 231 y 233.)
- [30] Elisa Drelie Gelasca: *Full-Reference Objective Quality Metrics for Video Watermarking, Video Segmentation and 3D Model Watermarking*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2005. (Citado en la página 184.)
- [31] Emrullah Durucan and Touradj Ebrahimi: *Change detection and background extraction by linear algebra*. *Proceedings of the IEEE*, 89(10):1368–1381, October 2001. (Citado en las páginas 58, 184 y 228.)
- [32] DB Elliott, MA Hurst, and J Weatherill: *Comparing clinical tests of visual function in cataract with the patients perceived visual disability*. *Eye*, 4:712–717, 1990. (Citado en la página 92.)
- [33] Angelika Erhardt-Ferron: *Theory and applications of digital image processing*. Textbook, Hochschule für Technik und Wirtschaft, 2000. (Citado en la página 22.)
- [34] M. Esteve, C.E. Palau, and T. Catarci: *A flexible video streaming system for urban traffic control*. *IEEE Multimedia*, 13(1):78–83, january–march 2006. <http://ieeexplore.ieee.org/search/wrapper.jsp?arnumber=1580436>. (Citado en la página 4.)
- [35] European Commission: *Work programme for the ict - information and communication technologies theme of the vii framework programme*. Website, June 2007. <ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2007-08.en.pdf>. (Citado en la página 10.)
- [36] F.A.M.: *Cibersoldado Español*. *Fuerzas Armadas del Mundo*, 05(5):10, Mayo 2005. (Citado en las páginas 11 y 54.)

- [37] Nicolás L. Fernández-García, A. Carmona-Poyato, Rafael Medina-Carnicer, and Francisco J. Madrid-Cuevas: *Automatic generation of consensus ground truth for the comparison of edge detection techniques*. *Image and Vision Computing*, 26(4):496–511, April 2008. (Citado en la página 53.)
- [38] Jerry R. Fram and Edward S. Deutsch: *On the quantitative evaluation of edge detection schemes and their comparison with human performance*. *IEEE Transactions on Computers*, C-24(6):616–628, June 1975. (Citado en la página 53.)
- [39] W. Frei and C.C. Chen: *Fast boundary detection: A generalization and a new algorithm*. *IEEE Transactions on Computers*, C-26(10):988–998, 1977. (Citado en la página 38.)
- [40] NA Frost, JM Sparrow, and JS Durant: *Development of a questionnaire for measurement of vision-related quality of life*. *Ophthalmic Epidemiol*, 5:185–210, 1998. (Citado en la página 95.)
- [41] Arturo Gallego Segador y Roberto Espejo Mohedano: *Fundamentos de Estadística*. Copisterías Don Folio S.L., ISBN: 84-930561-8-9, 2001. (Citado en las páginas 23, 34 y 58.)
- [42] Rafael C. Gonzalez and Richard E. Woods: *Digital Image Processing*. ISBN: 0-201-18075-8. Prentice-Hall, Inc., Upper Saddle River (NJ), USA, 2nd edition, 2002. (Citado en las páginas 22, 23, 33 y 120.)
- [43] José María González Linares: *Detección automática de objetos deformables*. Tesis de Doctorado, Universidad de Málaga, 2000. (Citado en la página 41.)
- [44] I. E. Gordon: *Theories of Visual Perception*. Wiley, 1997. (Citado en la página 120.)
- [45] Frank Grubb: *Procedures for detecting outlying observations in samples*. *Technometrics*, 11(1):1–21, Feb. 1969. (Citado en la página 197.)
- [46] Bob E. Hayes: *Cómo medir la satisfacción del cliente. Diseño de encuestas, uso y métodos de análisis estadístico*. Oxford University Press, México, 2ª edición, 1999. (Citado en la página 76.)
- [47] Michael D. Heath: *A robust visual method for assessing the relative performance of edge-detection algorithms*. Master's thesis, Graduate School. University of South Florida. Tampa, Florida, September 1996. ftp://figment.csee.usf.edu/pub/Edge_Comparison/Edge_Comparison_thesis.ps.gz. (Citado en las páginas 60, 118, 119, 148, 149, 184, 188 y 228.)
- [48] Michael D. Heath, Sudeep Sarkar, Thomas Sanocki, and Kevin W. Bowyer: *Comparison of edge detectors: a methodology and initial study*. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 143–148, San Francisco, USA, June 1996. (Citado en la página 118.)

- [49] Michael D. Heath, Sudeep Sarkar, Thomas Sanocki, and Kevin W. Bowyer: *A robust visual method for assessing the relative performance of edge-detection algorithms*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(12):1338–1359, December 1997. (Citado en las páginas [xix](#), [61](#) y [118](#).)
- [50] Michael D. Heath, Sudeep Sarkar, Thomas Sanocki, and Kevin W. Bowyer: *Comparison of edge detectors: A methodology and initial study*. Computer Vision and Image Understanding, 69(1):38–54, January 1998. (Citado en las páginas [60](#), [118](#), [184](#) y [228](#).)
- [51] A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. Goldgof, and Kevin W. Bowyer: *Range image segmentation: The user's dilemma*. In *International Symposium on Computer Vision*, pages 323–328, 1995. (Citado en la página [53](#).)
- [52] ITU-R: *Methodology for the subjective assessment of the quality of television pictures*. Recommendation BT.500–6, 1994. (Citado en la página [56](#).)
- [53] ITU-T: *Methods for subjective determination of transmissions quality*. Recommendation P.800, 1996. (Citado en las páginas [7](#) y [47](#).)
- [54] ITU-T: *Methods for subjective determination of video quality for multi-media applications*. Recommendation P.910, August 1996. (Citado en las páginas [47](#) y [184](#).)
- [55] ITU-T: *Principles of a reference impairment system for video*. Recommendation P.930, August 1996. (Citado en las páginas [56](#) y [184](#).)
- [56] Bernd Jähne: *Digital Image Processing*. ISBN: 3-540-67754-2. Springer-Verlag, Berlin Heidelberg, 5th edition, 2002. (Citado en las páginas [23](#) y [120](#).)
- [57] S.M. Jin, I.B. Lee, and J.M. Han: *Context-based pixelization model for the artificial retina using saliency map and skin color detection algorithm*. In *Human Vision and Electronic Imaging XIII - Proc. SPIE*, volume 6806, San Jose, CA, USA, January 2008. SPIE Press. (Citado en la página [4](#).)
- [58] Michel Jourlin and Jean Charles Pinoli: *Logarithmic image processing*. Acta Stereologica, 6:651–656, 1987. (Citado en la página [24](#).)
- [59] Michel Jourlin and Jean Charles Pinoli: *A model for logarithmic image-processing*. Journal of Microscopy, 149(1):21–35, January 1988. (Citado en la página [24](#).)
- [60] W Kosnik, L Winslow, and D Kline: *Visual changes in daily life throughout adulthood*. Journal of Gerontology, 43:63–70, 1988. (Citado en las páginas [91](#) y [92](#).)
- [61] Paula Lagares Barreiro y Justo Puerto Albandoz: *Población y Muestra. Técnicas de muestreos*, 2001. http://optimierung.mathematik.uni-kl.de/mamausch/index.php?content=ver_texte. (Citado en las páginas [72](#), [73](#) y [74](#).)

- [62] Jaron Lanier: *Brief biography of Jaron Lanier*. Webpage. <http://www.jaronlanier.com/general.html>. (Citado en la página 11.)
- [63] Layar Inc.: *Augmented reality – layar reality browser*, September 2010. <http://www.layar.com>. (Citado en la página 13.)
- [64] Gang Luo and Eli Peli: *Kinematics of visual search by tunnel vision patients with augmented vision see-through HMD*. *SID Digest*, ?:1578–1581, 2004. (Citado en las páginas 9 y 10.)
- [65] A. Mack and I. Rock: *Unattentional Blindness*. MIT Press/Bradford Books series in Cognitive Psychology, 1998. (Citado en la página 10.)
- [66] CM Mangione, RS Phillips, and JM Seddon: *Development of the activities of daily vision scale. a measure of visual functional status*. *Medical Care*, 30:1111–1126, 1992. (Citado en la página 94.)
- [67] David Marr and E Hildreth: *Theory of Edge-Detection*. *Proceedings of the Royal Society of London Series B – Biological Sciences*, 207(1167):187–217, 1980. (Citado en la página 40.)
- [68] Robert W. Massof and Gary S. Rubin: *Visual function assesment questionnaires*. *Survey of Ophthalmology*, 45(6):531–548, June 2001. (Citado en las páginas 89, 90, 94 y 97.)
- [69] Margaret W. Matlin y Hugh J. Foley: *Sensación y Percepción*. Prentice–Hall, Inc., Mexico D.F., 3ª edición, 1996. ISBN: 968–880–677–3. (Citado en las páginas 51 y 52.)
- [70] Kenneth O. McGraw and S.P. Wong: *Forming inferences about some intraclass correlation coefficients*. *Psychological Methods*, 1(1):30–46, 1996. (Citado en la página 86.)
- [71] Kevin McKoen, Raquel Navarro Prieto, Benoit Duc, Emrullah Durucan, Francesco Ziliani, and Touradj Ebrahimi: *Evaluation of video segmentation methods for surveillance applications*. In *Eusipco, X European Signal Processing Conference*, volume 2, pages 1–4, September 4–8 2000. (Citado en las páginas 54, 55, 56 y 58.)
- [72] V. Mezaris, I. Kompatsiaris, and M. G. Strintzis: *Still image objective segmentation evaluation using ground truth*. In B. Kovar, J. Prikryl, and M. Vlcek (editors): *5th COST 276 Workshop*, pages 9–14, 2003. (Citado en la página 53.)
- [73] V. Morellas, I. Pavlidis, and P. Tsiamyrtzis: *Deter: Detection of events for threat evaluation and recognition*. *Machine Vision and Applications*, 15(1):29–45, October 2003. (Citado en la página 4.)
- [74] U. Neisser and R. Becklen: *Selective looking: Attending to visually specified events*. *Cognitive Psychology*, 7:480–494, 1975. (Citado en la página 10.)
- [75] OMS: *Magnitude and causes of visual impairment*. Website, November 2004. <http://www.who.int/mediacentre/factsheets/fs282/en/index.html>, Fact Sheet N° 282. (Citado en las páginas 8 y 9.)

- [76] A. N. Oppenheim: *Questionnaire Design, Interviewing and Attitude Measurement*. Continuum, London and New York, new edition, 2000. (Citado en las páginas 64, 76 y 118.)
- [77] Alan V. Oppenheim: *Speech analysis–synthesis system based on homomorphic filtering*. *Journal of Acoustic Society of America*, 45(2):458–465, February 1969. http://www.rle.mit.edu/dspg/documents/speechanaly_1969.pdf. (Citado en las páginas 6, 23 y 29.)
- [78] Alan V. Oppenheim, Ronald W. Schafer, and Thomas G. Stockham JR.: *Nonlinear filtering of multiplied and convolved signals*. *IEEE Transactions on Audio and Electroacoustics*, AU-16(3):437–466, September 1968. http://www.rle.mit.edu/dspg/documents/nonlinearfiltering_1968.pdf. (Citado en las páginas 6, 22, 23 y 29.)
- [79] Rafael Palomar, José M. Palomares, , José M. Castillo, , Joaquín Olivares, , and Juan Gómez-Luna: *Parallelizing and optimizing LIP–Canny using NVIDIA CUDA*. In N. García-Pedrajas (editor): *Proceedings of the IEA/AIE 2010*, volume 6098 – Part III of *Lecture Notes on Artificial Intelligence, LNAI*, pages 389–398, Córdoba (España), Junio 2010. (Citado en la página 233.)
- [80] José M. Palomares, Jesús González, and Eduardo Ros: *Designing a fast convolution under the LIP paradigm applied to edge detection*. In Sameer Singh et al. (editor): *Proceedings of the III International Conference on Advances in Pattern Recognition ICAPR 2005*, volume 3 of *Lecture Notes in Computer Science – LNCS 3687*, pages 560–569, Bath, UK., 22–25, August 2005. (Citado en las páginas 106 y 233.)
- [81] José M. Palomares, Jesús González y Eduardo Ros: *Detección de Bordes en Imágenes con Sombras mediante LIP–Canny*. En *AERFAI 2005*, 2005. (Citado en la página 233.)
- [82] José M. Palomares, Jesús González, Eduardo Ros, and Alberto Prieto: *General logarithmic image processing convolution*. *IEEE Transactions on Image Processing*, 15(11):3602–3608, November 2006. ISSN: 1057–7149. (Citado en las páginas 106 y 233.)
- [83] Eli Peli, , and Fernando Vargas-Martin: *In-the-spectacle–lens tele-scopic device*. *Journal of Biomedical Optics*, 13(3):1–11, may/june 2008. (Citado en la página 4.)
- [84] Eli Peli: *Vision multiplexing: an engineering approach to vision rehabilitation device development*. *Optometry and Vision Science*, 78(5):304–315, May 2001. (Citado en la página 4.)
- [85] Jean Charles Pinoli: *A general comparative study of the multiplicative homomorphic, log–ratio and logarithmic image processing approaches*. *Signal Processing*, 58(1):11–45, April 1997. (Citado en las páginas 28, 29 y 30.)
- [86] Jean Charles Pinoli: *The logarithmic image processing model: Connections with human brightness perception and contrast estimators*.

- Journal of Mathematical Imaging and Vision, 7(4):341–358, 1997. (Citado en las páginas 24, 28, 30 y 32.)
- [87] J.M.S. Prewitt: *Picture Processing and Psychopictorics*, chapter Object Enhancement and Extraction. Academic Press, New York, 1970. (Citado en la página 37.)
- [88] Richard J. Radke, Srinivas Andra, Omar Al-Kafahi, and Badrinath Roysam: *Image change detection algorithms: a systematic survey*. IEEE Transactions on Image Processing, 14(3):294–307, March 2005. (Citado en la página 53.)
- [89] R.A.E. (editor): *Diccionario de la Lengua Española*, volumen 1. Real Academia Española de la Lengua, 22^a edición, 2002. (Citado en la página 65.)
- [90] Clarence E. Rash (editor): *Helmet Mounted Displays: Design Issues for Rotary-wing Aircraft*. SPIE Press, 2001. <http://books.google.es/books?id=fl60pXvwmLEC>. (Citado en la página 11.)
- [91] Clarence E. Rash: *A 25-year retrospective review of visual complaints and illusions associated with a monocular helmet-mounted display*. Displays, 29(2):70–80, March 2008. (Citado en la página 11.)
- [92] Georg Szalai (Reuters): *Video game industry growth still strong: study*, June 2007. <http://www.reuters.com/article/idUSN2132172920070621>. (Citado en la página 13.)
- [93] L.G. Roberts: *Optical and Electro-Optical Information Processing*, chapter Machine Perception of Three-Dimensional Solids, pages 159–197. MIT Press, Cambridge, Massachusetts, USA, 1965. (Citado en la página 37.)
- [94] Eduardo Ros, Javier Díaz, Sonia Mota, Fernando Vargas-Martin, and M.D. Peláez-Coca: *Real time image processing on a portable aid device for low vision patients*. In K. Bertels, J.M.P. Cardoso, and S. Vassiliadis (editors): *ARC 2006*, LNCS 3985, pages 158–163. Springer-Verlag, Berlin Heidelberg, 2006. (Citado en la página 4.)
- [95] PW Russell, R Sekuler, and C Fetkenhour: *Visual function after pan-retinal photocoagulation: a survey*. Diabetes Care, 8:57–63, 1985. (Citado en la página 91.)
- [96] H. Scharr and J. Weickert: *An anisotropic diffusion algorithm with optimized rotation invariance*, 2000. <http://citeseer.ist.psu.edu/scharr00anisotropic.html>. (Citado en la página 38.)
- [97] S.S. Shapiro and M.B. Wilk: *An analysis of variance test for normality (complete samples)*. Biometrika, 52(3–4):591–611, 1965. (Citado en la página 197.)
- [98] Kim Sheehan: *E-mail survey response rates: A review*. Journal of Computer-Mediated Communication, 6(2):1–13, January 2001. (Citado en la página 190.)

- [99] Patrick E. Shrouf and Joseph L. Fleiss: *Intraclass correlations: Uses in assessing rater reliability*. *Psychological Bulletin*, 86(2):420–428, 1979. (Citado en las páginas 85 y 210.)
- [100] Haim Shvayster and Shmuel Peleg: *Pictures as elements in a vector space*. In *Proceedings of the 1983 IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 442–446, Washington, USA., June 1983. (Citado en la página 29.)
- [101] Haim Shvayster and Shmuel Peleg: *Inversion of picture operators*. *Pattern Recognition Letters*, 5(1):49–61, January 1987. (Citado en la página 29.)
- [102] ME Sloane, K Ball, and C Owsley: *The visual activities questionnaire: developing an instrument for assessing problems in everyday visual tasks*. Technical Report 1, Optical Society of America: OSA Tech Dig Noninvasive Assess Vis Sys, Washington DC., USA, 1992. pp. 26–29. (Citado en la página 93.)
- [103] Irwin Sobel: *Camera Models and Machine Perception*. Ph.d. thesis, Stanford University, 1970. (Citado en la página 37.)
- [104] Irwin Sobel and G. Feldman: *A 3x3 isotropic gradient operator for image processing*. Talk at the Stanford Artificial Project, 1968. (Citado en la página 37.)
- [105] Mobilizy Mobile Software: *Wikitude*, September 2010. <http://www.wikitude.org>. (Citado en la página 13.)
- [106] EP Steinberg, JM Tielsch, and OD Schein: *The vf-14. an index of functional impairment in patients with cataract*. *Arch Ophthalmol*, 112:630–638, 1994. (Citado en la página 95.)
- [107] JP Szlyk, A Arditi, and P Coffey-Bucci: *Self-report in functional assessment of low vision*. *Journal of Visually Impaired and Blind*, 84:61–66, 1990. (Citado en la página 92.)
- [108] S. Thorpe, A. Delorme, and R. Van Rullen: *Spike-Based Strategies for Rapid Processing*. *Neural Networks*, 14(6-7):715–725, July–September 2001. Special Issue. (Citado en la página 5.)
- [109] S. Thorpe, A. Delorme, R. Van Rullen, and W. Paquier: *Reverse Engineering of the Visual System using Networks of Spiking Neurons*. In *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS–2000*, volume 4, pages 405–408, Geneva, Switzerland, April 2000.
- [110] S. Thorpe, D. Fize, and C. Marlot: *Speed of Processing in the Human Visual System*. *Nature*, 6582(381):520–522, June 1996. (Citado en la página 5.)
- [111] M. Valera and S.A. Velastin: *Intelligent distributed surveillance systems: a review*. *IEE Proceedings - Vision, Image and Signal Processing*, 152(2):192–204, April 2005. (Citado en la página 4.)

- [112] Fernando Vargas-Martin: *A free-cost visual field expander for peripheral vision loss*. *Journal of Vision*, 3(12):40–40, December 2003. <http://journalofvision.org/3/12/40/>. (Citado en la página 54.)
- [113] Fernando Vargas-Martin and Eli Peli: *Augmented-view for restricted visual field: Multiple device implementations*. *Optometry and Vision Science*, 79(11):715–723, 2002. (Citado en las páginas 4 y 10.)
- [114] D. Vernon: *Machine Vision — Automated Visual Inspection and Robot Vision*. Prentice-Hall, Inc., 1991. (Citado en la página 165.)
- [115] David S. Walonick: *Excerpts from: "Survival Statistics"*. StatPac, Inc., Bloomington, MN, USA, 2004. ISBN: 0-918733-11-1. (Citado en las páginas 187, 189 y 190.)
- [116] Stefan Winkler: *Digital Video Quality — Vision Models and Metrics*. John Wiley & Sons, Inc., 2005. <http://books.google.com/books?id=NDNfMaht37cC>, ISBN 0-470-02404-6. (Citado en la página 54.)
- [117] Stefan Winkler: *Video quality and beyond*. In *Proceedings of the European Signal Processing Conference*, Poznan, Poland, september 3–7 2007. invited paper. (Citado en las páginas 48 y 54.)
- [118] Yitzhak Yitzhaky and Eli Peli: *A method for objective edge detection evaluation and detector parameter selection*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):1027–1033, August 2003. (Citado en la página 53.)
- [119] P. Zuidema, J.J. Koenderink, and M.A. Bouman: *A mechanistic approach to threshold behavior of the visual system*. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:923–933, 1983. (Citado en la página 30.)

DECLARACIÓN

Declaro que he desarrollado esta Tesis Doctoral, bajo la supervisión de los directores de Tesis y, por tanto, asumo la autoría de todo lo descrito en el presente documento.

También declaro que el trabajo realizado es inédito (salvo por las publicaciones científicas obtenidas y mencionadas en el apartado de PRODUCCIÓN CIENTÍFICA del Capítulo 12) y no es plagio ni total ni parcial de ninguna otra investigación realizada por otras personas.

Afirmo que todos los datos expuestos en esta investigación no han sido falseados y que cualquier error que pudiera existir en el documento, no ha sido introducido conscientemente.

Granada, noviembre 2010

José Manuel Palomares Muñoz

