



Universidad de Granada

FUSIÓN BAYESIANA DE PARES DE IMÁGENES CON DIFERENTE EXPOSICIÓN

MEMORIA QUE PRESENTA
MIGUEL TALLÓN BENÍTEZ
PARA OPTAR AL GRADO DE DOCTOR EN INFORMÁTICA
NOVIEMBRE 2012

DIRECTORES
RAFAEL MOLINA SORIANO
JAVIER MATEOS DELGADO

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
E INTELIGENCIA ARTIFICIAL

E.T.S. INGENIERÍAS INFORMÁTICA Y TELECOMUNICACIÓN UNIVERSIDAD DE GRANADA

Editor: Editorial de la Universidad de Granada
Autor: Miguel Tallón Benítez
D.L.: GR 854-2013
ISBN: 978-84-9028-452-0

La memoria titulada **Fusión Bayesiana de pares de imágenes con diferente exposición**, que presenta D. Miguel Tallón Benítez para optar al grado de DOCTOR, ha sido realizada en el Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada bajo la dirección de los Doctores D. Rafael Molina Soriano y D. Javier Mateos Delgado.

El doctorando D. Miguel Tallón Benítez y los directores de la tesis D. Rafael Molina Soriano y D. Javier Mateos Delgado garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y, hasta donde nuestro conocimiento alcanza, en la realización del trabajo se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, noviembre de 2012

El doctorando

Los directores

Miguel Tallón Benítez

Rafael Molina Soriano

Javier Mateos Delgado

AGRADECIMIENTOS

*El hombre nunca sabe de lo que es capaz
hasta que lo intenta.*

Charles Dickens (1812-1870).
Novelista inglés.

Quiero agradecer a los directores de esta tesis, Rafael Molina Soriano y Javier Mateos Delgado, por las grandes dosis de paciencia que han dedicado en mi formación como investigador además de brindarme la oportunidad de trabajar con ellos y colaborar con expertos de otros países. En especial con el Profesor Aggelos K. Katsaggelos y S. Derin Babacan. Sin ellos no habría sido posible la realización de este trabajo de investigación.

A todos mis amigos y compañeros que, de una forma u otra, han contribuido al desarrollo de esta Tesis.

Para terminar quiero dedicar este trabajo a mis seres más queridos, mi novia Clara y mis padres y agradecerles especialmente el apoyo recibido en todo momento. También al resto de mi familia y a Nancy Lu y Debbie que me acogieron de manera excepcional en las sucesivas estancias.

Gracias a todos.

A todo el que ha contribuido a que salga adelante el trabajo de investigación.

**FUSIÓN BAYESIANA DE PARES DE IMÁGENES CON DIFERENTE
EXPOSICIÓN**

MIGUEL TALLÓN BENÍTEZ

Índice general

Summary and conclusions	1
Summary	1
Goals and outline of the thesis	3
Conclusions	5
Future works	6
1. Introducción	9
1.1. Introducción a la combinación de imágenes	9
1.2. Proceso de formación de la imagen	13
1.3. Objetivos y estructura de la memoria	16
2. Restauración de pares de imágenes: Revisión histórica	19
2.1. Métodos de restauración basados en una única observación	24
2.2. Métodos basados en la combinación de observaciones	27
2.2.1. Métodos basados en PSF espacialmente invariante	28
2.2.2. Métodos basados en PSF espacialmente variante	28
2.3. Formulación Bayesiana jerárquica	31
2.3.1. Descripción del paradigma bayesiano para exposición dual	31
2.3.2. Soluciones al problema de la restauración	32
3. Calibración de las imágenes	37
3.1. Calibración fotométrica	38
3.1.1. Ajuste de la ganancia	39
3.1.2. Ecuación de histograma mediante histograma de imagen borrosa	41
3.2. Calibración geométrica	42

3.2.1. Registrado por emparejamiento de bloques (<i>Block Matching</i>)	43
3.2.2. Registrado mediante Homografía (Surf+RANSAC)	44
4. Deconvolución usando un núcleo espacialmente invariante	51
4.1. Introducción	51
4.2. Formulación del problema	52
4.3. Modelo Bayesiano jerárquico	54
4.3.1. Primera etapa: Modelos a priori	54
4.3.2. Primera etapa: Modelo de observación	56
4.3.3. Segunda etapa: Distribuciones sobre los hiperparámetros	57
4.4. Inferencia Bayesiana Variacional	58
4.5. Cálculo de las aproximaciones de la distribución a posteriori	62
4.6. Experimentos	66
4.6.1. Experimentos sintéticos	67
4.6.2. Experimentos reales	70
5. Métodos de deconvolución basados en núcleo espacialmente variante	77
5.1. Introducción	77
5.2. Preprocesamiento	79
5.3. Estimación del núcleo de emborronamiento	81
5.4. Deconvolución rápida	82
5.4.1. Etapa de reparación de la PSF	83
5.4.2. Deconvolución	84
5.5. Deconvolución basada en combinación de modelos de imagen	89
5.5.1. Deconvolución	90
5.6. Fusión de bloques	97
5.7. Experimentos	98
5.7.1. Experimentos sintéticos	99
5.7.2. Experimentos reales	106
6. Conclusiones y trabajo futuro	123

Índice de figuras

1. Photographs in dim conditions: (a) correct exposure image but blurred ($f/5.6$, $1/10s$, ISO 100); (b) Sharp image but dark and noisy ($f/5.6$, $1/160s$, ISO 100). . . .	2
1.1. (a) Partes de la cámara y (b) flujo que sigue la luz hasta reproducirse en el visor o realizar la captación por parte del sensor. Imágenes cortesía de http://www.fotografiapractica.com y http://www.textoscientificos.com , respectivamente	10
1.2. Triángulo de exposición. Los tres factores que influyen en la exposición y los efectos que pueden ocasionar.	11
1.3. Imágenes tomadas en situaciones de baja luminosidad: (a) imagen correctamente expuesta pero borrosa ($f/5.6$, $1/10s$, ISO 100); (b) imagen nítida pero ruidosa ($f/5.6$, $1/160s$, ISO 100).	13
1.4. Modelo de formación de imagen y revelado digital.	14
2.1. Tipos de ruido (de izquierda a derecha): Ruido de patrón fijo, ruido aleatorio y <i>banding noise</i> . Fuente: http://www.cambridgeincolour.com/	22
3.1. (a) Imagen borrosa e histograma. (b) Imagen de corta exposición con su histograma. (c) Resultado de aplicar método directo estimación de ganancia ($\lambda = 33,3$), debajo su histograma. (d) Resultado del calibrado fotométrico de aplicar método de mínimos cuadrados para la estimación de ganancia ($\lambda = 24,3$), debajo su histograma.	40

3.2. (a) Imagen borrosa, histograma e histograma acumulado. (b) Imagen de corta exposición, su histograma y el histograma acumulado. (c) Resultado de aplicar método de ecualización de histograma con histograma de la imagen borrosa (a), sobre la imagen oscura (b), debajo se muestra su histograma y el histograma acumulado.	47
3.3. Transformación obtenida para cada uno de los «bins», en el eje de abscisas, con su correspondiente valor de gris de salida en ordenadas.	48
3.4. (a) imagen borrosa sin calibrar, (b) Imagen borrosa calibrada mediante emparejamiento de bloques, con un desplazamiento calculado de $(-2, -2)$ en horizontal y vertical, respectivamente. (c) imagen ruidosa calibrada fotométricamente.	48
3.5. (a) imágenes de entrada con descriptores Surf. Cada línea une puntos en correspondencias. (b) Imagen borrosa rectificada (con la homografía aplicada). (c) Imagen borrosa (enmarcada) sin calibrado geométrico. (d) Imagen borrosa calibrada y enmarcada. (e) Imagen ruidosa enmarcada y calibrada fotométricamente para comparación visual del registrado.	49
4.1. Modelo jerárquico propuesto.	55
4.2. (a) Imagen original. (b) Observación ruidosa simulando imagen de corta exposición tras la calibración. (c) Método de supresión de ruido en [1] aplicado sobre la imagen ruidosa.	67
4.3. Imágenes borrosas simulando la imagen de exposición correcta en entornos de baja luminosidad. La PSF utilizada para generar las observaciones se muestra debajo de cada una de las imágenes. Nótese que los valores de las funciones de emborronamiento se han llevado al rango $[0,255]$ por razones de visualización.	68
4.4. Resultados de aplicar nuestro método seleccionando el λ óptimo sobre las imágenes borrosas de la Figura 4.3 y ruidosa 4.2b, debajo se muestra la PSF estimada en cada experimento.	69
4.5. Evolución de la PSNR en función de λ_1 para los resultados de la Figura 4.4.	70
4.6. Resultados de aplicar el método [2] a las imágenes borrosas de la Figura 4.3 y 4.2b, debajo se muestra la PSF estimada.	71
4.7. Resultados de aplicar el método deconvblind usando la PSF estimada por el método propuesto a las imágenes borrosas de la Figura 4.3 y 4.2b.	72

4.8. Experimento real 1: (a) observación borrosa, (b) ruidosa después de la calibración (c) restauración y PSF obtenidas por la aplicación del método en [2], (d) restauración y PSF obtenidas utilizando el método propuesto con $\lambda_1 = 0,8$	74
4.9. Experimento real 2: (a) observación borrosa, (b) ruidosa después de la calibración (c) restauración y PSF obtenidas por la aplicación del método en [2], (d) restauración y PSF obtenidas utilizando el método propuesto con $\lambda_1 = 0,8$	75
5.1. Ventana de Hann utilizada en la etapa de fusión de bloques.	98
5.2. (a) Imagen original, (b) imagen subexpuesta simulada, calibrada fotométricamente con la imagen borrosa mostrada en la Figura 5.3c.	99
5.3. Imágenes borrosas simulando larga exposición. Las PSFs utilizadas para generar cada una de las imágenes borrosas se muestran en la segunda fila. Nótese que los valores de las funciones de emborronamiento se han llevado al rango [0,255] por razones de visualización.	100
5.4. Restauraciones resultado de aplicar el Algoritmo 3 (AREVDR) en las imágenes sintéticas de la Figura 5.3. En la fila inferior se muestran los núcleos de emborronamiento estimados.	101
5.5. Restauraciones resultado de aplicar el Algoritmo 6 (AREVDCM) en las imágenes sintéticas de la Figura 5.3. Fila superior: Usando el valor óptimo de λ . Fila central: Usando el valor estimado de λ . Fila inferior: Núcleos de emborronamiento estimados.	111
5.6. Restauraciones resultado del experimento sintético de la Figura 5.3c y 5.2b: (a) imagen con supresión de ruido usando el método [1], (b) deconvblind aplicando nuestro marco de bloques para emborronamiento espacialmente variante, (c) deconvblind asumiendo emborronamiento espacialmente invariante, (d) método en [2] (espar.), (e) método original en [2], (f) AREI, (g) AREVDR, (h) AREVDCM ($\lambda = 0$), (i) AREVDCM($\lambda = 1$), (j) AREVDCM.	112
5.7. PSFs estimados en los experimentos de la Figura 5.6: (a) PSF global estimada en deconvblind, (b) PSF global en [2], (c) PSF global en AREI, (d) PSFs reales, (e) PSF espacialmente variante estimada en deconvblind, (f) PSFs asumiendo emborronamiento espacialmente variante en [2], (g) PSFs estimadas en los métodos propuestos. Cada PSF individual de tamaño 15×15	113

5.8. Resultados del experimento sintético de la Figura 5.3b con solapamiento y configuración 7×5 bloques aplicado a AREVDR. (a) PSFs estimadas. (b) bloques donde se ha aplicado la etapa de corrección de emborronamiento. (c) PSFs corregidas. (d) Imagen restaurada con una PSNR de 25,6 dB.	113
5.9. Resultados del experimento sintético de la Figura 5.3b con solapamiento y configuración 7×5 bloques aplicado a AREVDCM: (a) Evolución de la PSNR para distintos valores del umbral t usando los valores estimados de λ . (b) PSF estimadas. (c) bloques con eliminación de ruido con $t = 1,5$. (d) Imagen restaurada con $t = 1,5$	114
5.10. Imágenes de prueba después del calibrado: Izquierda, imagen de larga exposición y derecha, imagen de corta exposición observada.	115
5.11. Resultados para el primer experimento: (a) imagen tras supresión de ruido usando [1], (b) restauración con el método en [2] (c) imagen restaurada con el método [3], (d) resultado de aplicar AREI, (e) resultado de aplicar AREVDR (f) resultado de aplicar AREVDCM.	116
5.12. Resultados para el segundo experimento: (a) imagen tras supresión de ruido usando [1], (b) restauración con el método en [2] (c) imagen restaurada con el método [3], (d) resultado de aplicar AREI, (e) resultado de aplicar AREVDR (f) resultado de aplicar AREVDCM.	117
5.13. Resultados para el tercer experimento: (a) imagen tras supresión de ruido usando [1], (b) restauración con el método en [2] (c) imagen restaurada con el método [4], (d) resultado de aplicar AREI, (e) resultado de aplicar AREVDR (f) resultado de aplicar AREVDCM.	118
5.14. Núcleos de emborronamiento estimados por el método [2] (primera fila) y AREI del Capítulo 4 (segunda fila) para cada experimento real. Los valores de las PSFs fueron normalizados al rango [0,255] por motivos de visualización. (a) y (d) Núcleo obtenido para las imágenes en las Figuras 5.10a y 5.10b, (b) y (e) PSF obtenida en el experimento de las Figuras 5.10c y 5.10d, (c) y (f) Núcleo de emborronamiento para las Figuras 5.10e y 5.10f.	119

- 5.15. Núcleos de emborronamiento estimados y corregidos junto a la estrategia de selección para cada uno de los experimentos reales ejecutados sobre el AREVDR. Los bloques en blanco representan parches en los que se aplicó eliminación de ruido o corrección de PSF. Los valores de las PSFs fueron normalizados al rango [0,255] por motivos de visualización. Cada fila corresponde con los experimentos en las Figuras 5.10a y 5.10b, 5.10c y 5.10d, 5.10e y 5.10f, respectivamente: (a),(e),(i) Núcleos obtenidos en cada experimento. (b),(f),(j) Bloques en los que se aplica la etapa de corrección. (c),(g),(k) PSFs corregidas. (d),(h),(l) Estrategia de selección donde se aplica el método en [1]. 120
- 5.16. Núcleos de emborronamiento estimados y estrategia de selección para cada uno de los experimentos reales aplicando el AREVDCM. Los bloques en blanco representan parches en los que se aplicó eliminación de ruido. Los valores de las PSFs fueron normalizados al rango [0,255] por motivos de visualización. (a) Núcleos obtenidos para las imágenes en las Figuras. 5.10a y 5.10b. (b) Estrategia de selección para la restauración de la Figura 5.11f. (c) PSFs obtenidas en el experimento de las Figuras 5.10c y 5.10d. (d) Estrategia de selección para el experimento de la Figura 5.12f. (e) Núcleos de emborronamiento para las Figuras 5.10e y 5.10f. (f) Estrategia de selección en la restauración de la Figura 5.13f. 121

Índice de cuadros

4.1. Métrica PSNR (Peak Signal to Noise Ratio) para las imágenes de los experimento sintéticos	73
5.1. Métricas PSNR y SSIM (Peak Signal to Noise Ratio y Structural Similarity) para las imágenes de los experimento sintéticos de la Figura 5.3.	102
5.2. Comparación en PSNR para el conjunto de PSFs estimadas en los experimentos sintéticos.	103
5.3. Valor de λ óptimo a la izquierda y el estimado a la derecha para cada uno de los bloques de la Figura 5.3c. Para la ejecución de AREVDCM (Alg. 6)	103
5.4. Tiempo de ejecución (en segundos) de cada uno de los métodos comparados para los experimentos en la imagen de la Figura 5.3c.	103

Summary and conclusions

*In critical moments even the very powerful have
need of the weakest.*

Aesop (620-564 BC).

Greek's fabulist.

A fin de cumplir con lo establecido en la normativa vigente de regulación de las enseñanzas oficiales de Doctorado y del título de Doctor por la Universidad de Granada aprobadas por Consejo de Gobierno de la Universidad de Granada en su sesión del 2 de Mayo del 2012. Este capítulo presenta un resumen de la introducción (Capítulo 1) y conclusiones principales (Capítulo 6) de esta tesis en lengua inglesa.

Summary

Image restoration began in the 60s with the American and Soviet space programs to enhance the images captured with devices with limitations, such as vibrations, optical aberrations, misalignment in capturing devices or lack of lighting. The images exhibited two main types of degradation: blur and noise.

The goal of digital image restoration is to recover the original image from the blurred observations. In many real applications, since there is more unknowns than observed data. For instance, from a single blurred image, the blurring function that caused the degradation has to be estimated and utilized to deconvolve the degraded image in a process named blind deconvolution (BD). Incorporating additional observations to the process helps in several ways, such as producing a more precise estimation of the blurring function, and also by better conditioning the deconvolution problem. However it also brings new challenges, such as the correction of the possible misalignment between the images.

Currently, Digital Single-Lens Reflex (DSLR) cameras incorporate mechanisms to assist the user to obtain a correctly exposed image under proper lighting conditions. Unfortunately, to acquire satisfying pictures in dim environments is, so far, a challenging task.

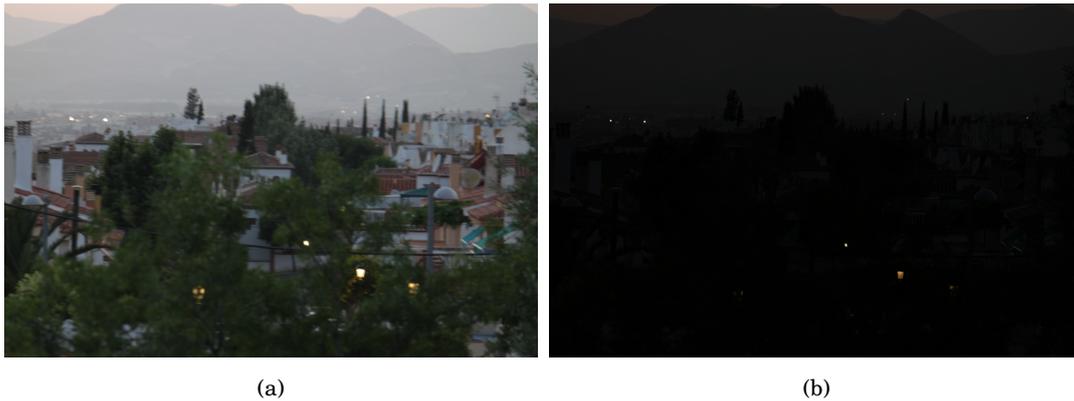


Figura 1: Photographs in dim conditions: (a) correct exposure image but blurred ($f/5.6$, $1/10s$, ISO 100); (b) Sharp image but dark and noisy ($f/5.6$, $1/160s$, ISO 100).

The choice of an appropriate configuration of aperture, exposure time and ISO sensitivity, in conjunction with mechanisms such as flash, the use of tripod or image stabilizer systems that incorporate current DSLR cameras are useful to capture images under certain illumination conditions. Unfortunately, there are many situations where the lighting is not adequate and these mechanisms may not be sufficient to prevent the degradation for appearing (see Figure 1). Moreover, flash could be insufficient to photograph wide scenes or those far away. Furthermore, the tripod could be a cumbersome instrument and its use is conditioned by the supporting surface. Regarding the image stabilizer, this solution is useful against slight camera shake but insufficient in case of severe blur. The image capture related problems are accentuated when using a low-end/entry level camera model or mobile phone camera since these devices do not capture as much light as DSLR cameras, and often do not provide some of the mechanisms described above. In low light conditions, it is likely that none of the above mechanisms are sufficient to obtain a sharp image. It is specially, in those situations where the algorithms presented in this dissertation perfectly fit.

The objective of this thesis is to recover a sharp image in low light conditions by combining information from images with different exposures. In particular, we will demonstrate that a pair of images, consisting of a correctly exposed but blurred image and a sharp but very noisy

image due to the short exposure time, as shown in Fig. 1, is enough to recover an image free of blurring and noise, by the combination of both images in the pair.

Goals and outline of the thesis

The main goal of this dissertation is to obtain a crisp correctly exposed image by combining a pair of images taken in low-light environments. Typically, the image pair taken under dim conditions, is composed by a blurred but correctly exposed image and another sharp but underexposed image, with low signal-to-noise-ratio (SNR). This main goal is further divided into the following set of objectives:

- To calibrate the images geometrically and photometrically as a preprocessing stage for the restoration algorithm.
- To use the hierarchical Bayesian paradigm to model the problem due to its versatility and theoretical foundations.
- To develop methods and techniques for combining images with different exposures, minimizing the blur and noise in the restored image.
- To study the combination of different observation models and a priori distributions on the image.
- To estimate the hierarchical model parameters in an automatic way.
- To propose a solution to spatially variant blurring.
- To implement and validate the proposed models, by experimenting on real and synthetic images, and comparing with other methods.

In order to achieve those objectives, we propose the following outline: In Chapter 1 we present the problem of capturing images in low-light environments, enumerate the main problems and describe the camera mechanisms to alleviate some of them. Furthermore, the image formation process is presented and the goals of the thesis are enumerated.

The second chapter reviews the blind restoration literature from a single image and from image pairs. In addition, it presents a deep explanation of the types of degradation found in the images and proposes a taxonomy to organize the different image restoration techniques

from image pairs. Finally, we introduce the theoretical foundations that serve as basis of the methods that we will develop: the hierarchical Bayesian formulation.

The third chapter presents several solutions to photometrically and geometrically calibrate the image pair, as a preprocessing step to the application of any of the restoration algorithms developed.

Chapters 4 and 5 are the core of this dissertation. In these chapters, restoration algorithms for dual exposure are proposed. More specifically, in Chapter 4 a spatially invariant blur model is assumed, while in Chapter 5 a more realistic model is proposed, intended for images degraded by spatially variant blur. The hierarchical Bayesian paradigm was utilized to develop all the restoration algorithms proposed in this dissertation, providing a comprehensive mathematical model that allows us to represent our knowledge about the images and infer knowledge about the original image. In general, the hierarchical Bayesian paradigm consists of a series of models: observation, image and blur model, which allow us to add knowledge about the problem. Each model depends on a series of parameters that should be estimated jointly with the sharp image. We define probability distributions on these parameters to include our knowledge about the possible value of the parameter as well. This paradigm provides a robust framework to incorporate both, the knowledge and uncertainty of our models, in addition to the advantage of automatically estimate their associated parameters.

Then, in the last chapter, we present the conclusions and future work that has emerged during the research. The dissertation concludes with the bibliography.

The work presented in this thesis has been partially published in the following three international conferences and a JCR indexed journal:

- M. Tallón, J. Mateos, S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Combining observation models in dual exposure problems using the Kullback-Leibler divergence,” in *European Signal Processing Conference (EUSIPCO)*, 2010, pp. 323–327
- M. Tallón, J. Mateos, R. Molina, and A.K. Katsaggelos, “Image prior combination in space-variant blur deconvolution for the dual exposure problem,” in *7th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2011, pp. 408–413
- M. Tallón, J. Mateos, S.D. Babacan, R. Molina, and A.K. Katsaggelos, “Space-variant kernel deconvolution for dual exposure problem,” in *19th European Signal Processing Conference (EUSIPCO)*, 2011, pp. 1678–1682

- M.Tallón, J. Mateos, S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Space-variant blur deconvolution and denoising in the dual exposure problem,” *Information Fusion*, pp.–, 2012

Conclusions

In this dissertation, the problem of obtaining sharp images in low light has been considered. The wide range of scenarios in which may be useful attracted our interest. To solve this problem a number of preprocessing steps had to be carried out, before proceeding into the restoration process. As we progressed in the investigation, we found that the traditional methods that consider a spatially invariant blur are useful in specific situations but limited in the restoration of real scenes where moving objects at different depths are involved. To address these limitations we considered a spatially variant blur model more flexible which provided better results. Let us now to detail the major contributions and the conclusions of the work.

- A review of blind deconvolution methods is presented, for those using a single image as well as for those addressing the problem of dual exposure. In addition, a taxonomy of deconvolution methods from image pairs based on the type of blur has been proposed. We provide several reasons why it is advantageous to include the low-exposure image deconvolution process. Despite the problems that entails, such as the need of another shot, calibration issues, etc., this second observation allows to reduce the ill-posedness of the deconvolution problem, leading to an easier modeling, a more accurate estimate of blur kernel and the recovery of a sharper image.
- Several techniques for photometric and geometric calibration have been analyzed. A robust and automatic calibration method for image pairs with different exposure time was proposed.
- We studied the blind deconvolution methods that combine images with different exposure time, considering an invariant blur and, thus, an algorithm which combines observation models is proposed. The proposed method achieved better results than the competitors we compared to. Unfortunately, due to the assumed invariant structure of the blur, the proposed method did not provide satisfactory results in some real problems.

- Due to the limitations of the methods that consider an invariant blurring model, two restoration algorithms to remove spatially variant blur are proposed. The results demonstrate that the spatially variant framework optimally solves the dual exposure problem. There is, however, a tradeoff between the size of the block and the blur size used in the variant definition. Note that in case of severe blur a larger block size is needed to carry out the estimation. Thus, a smaller number of blocks may affect the potential of the proposed framework.
- The proposed methods were developed using a strong theoretical basis, and require little or none user supervision, with automatic estimation of the parameters of the algorithm. The problem is satisfactorily solved in both, the linear domain of the sensor and in the JPEG image domain. Nevertheless, the developed algorithms suffer from very high computational cost to be implemented in real time.
- The proposed methods are applicable not only to images taken with a mobile camera, where the effects of blur and noise are most dramatic due to the small dimensions of the sensor, but it can be extended to other scenarios such as medical imaging, video surveillance sequences, depth map super-resolution or high dynamic range (HDR) imaging.

Future works

- To optimize the Algorithm 3: fast spatially variant deconvolution so that it can be executed in mobile phones. Also, to develop a massively parallel implementation of algorithm using GPUs.
- The estimation of the parameter λ (which controls the contribution of each image model) as another unknown parameter within the Bayesian hierarchical model and to dynamically adjust the block size. To achieve the latter, segmentation methods may be used, which would increase the flexibility at the expense of increasing the computation cost.
- To develop a method capable of combining more than two images for different purposes. The first one, directly related to the objective of this thesis, is to combine images to obtain a sharp image. A second one would be to recover a high dynamic range image similar to the one obtained by using the HDR technique. In the case, our solution would avoid the

propagation of the blur of images with longer exposure time. We could also take advantage of the exposure bracketing technique, to combine multiple photographs at different exposures in a single shot.

Capítulo 1

Introducción

Una obra nunca se acaba, sino que llega al límite de las propias posibilidades.

Antonio López.
Pintor y escultor.

1.1. Introducción a la combinación de imágenes

La restauración de imágenes surge en los años 60 con motivo de la carrera espacial entre americanos y soviéticos para la obtención de imágenes mejoradas que fueron adquiridas con dispositivos con algunas limitaciones, tales como: vibraciones, aberraciones ópticas, desalineación de los aparatos de captación o la falta de luminosidad. Las imágenes captadas presentaban principalmente dos tipos de degradaciones, el emborronamiento y el ruido.

El objetivo de la restauración digital de imágenes es el de recuperar una imagen original a partir de las observaciones borrosas. **En muchas aplicaciones reales es preciso estimar la función de emborronamiento que produjo la degradación y deshacerla en un proceso conocido como deconvolución.** Incorporar observaciones adicionales al proceso ayuda en varios aspectos, tales como, la estimación de la función de emborronamiento y condiciona mejor la deconvolución. Sin embargo supone nuevos retos como la corrección de la posible desalineación entre las imágenes.

En la actualidad, las cámaras DSLR (*Digital Single-Lens Reflex*) como el modelo de la Figura 1.1a, incorporan mecanismos para ayudar al usuario en la obtención de una imagen correctamente expuesta en condiciones de luminosidad adecuadas. Desafortunadamente, en entornos de baja luminosidad es muy difícil obtener una buena fotografía. La Figura 1.1b muestra la

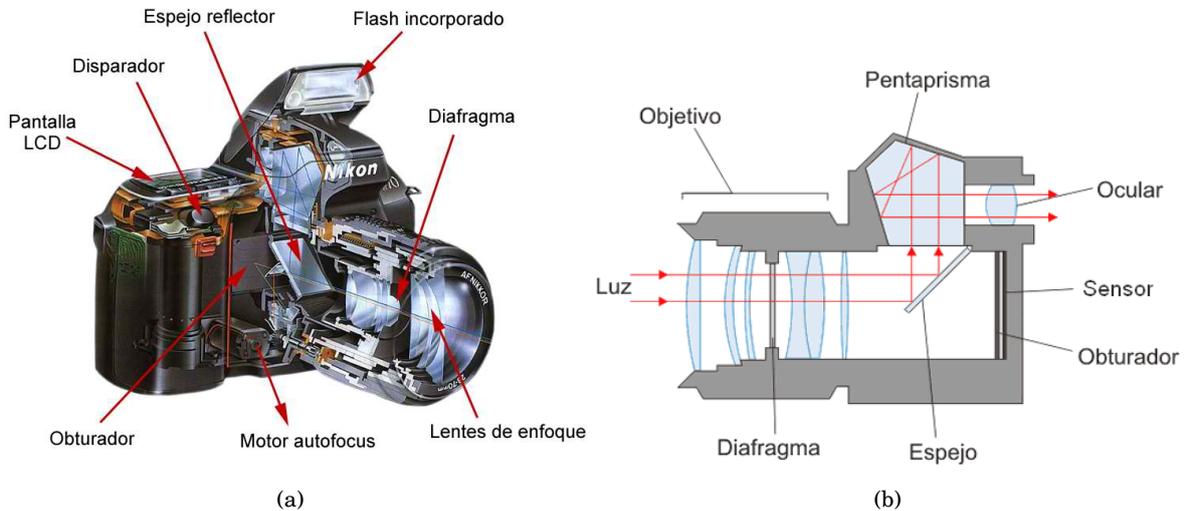


Figura 1.1: (a) Partes de la cámara y (b) flujo que sigue la luz hasta reproducirse en el visor o realizar la captación por parte del sensor. Imágenes cortesía de <http://www.fotografiapractica.com> y <http://www.textoscientificos.com>, respectivamente

trayectoria que sigue la luz a través del diafragma, atravesando una serie de lentes, hasta finalmente alcanzar el visor o el sensor de la cámara (según la posición del espejo). Cada uno de los píxeles del sensor posee un fotodiodo encargado de recoger y almacenar los fotones. Cuando la exposición concluye, la cámara evalúa los fotones de cada una de las celdas para formar la imagen. El proceso de formación de la imagen se describe en detalle en la siguiente sección. Nótese que la posición del espejo de la Figura 1.1b refleja la luz hacia el prisma pentagonal para reproducir la escena a la que enfocamos en el visor; la otra posición del espejo, levantada hacia el pentaprisma, permite que los fotones alcancen el sensor. A continuación explicamos los tres parámetros que gobiernan los mecanismos de la cámara que influyen en la exposición.

Primeramente y para entender la exposición nos valemos de un ejemplo sencillo como el de recoger agua de lluvia con un cubo [10], partiendo de que somos capaces de controlar tres factores: el ancho del cubo, la duración que permanece bajo la lluvia y la cantidad de agua que deseamos recoger. Debemos asegurarnos que la cantidad de agua recogida no sea muy poca (subexposición) pero tampoco demasiada (sobrexposición). La clave reside en la existencia de multitud de combinaciones de ancho, tiempo y cantidad que consigan el objetivo. Imaginemos, el caso en el que dejamos constante la cantidad de agua que deseamos recoger, este nivel podríamos conseguirlo con menos tiempo bajo la lluvia pero cogiendo un cubo más grande. Alter-

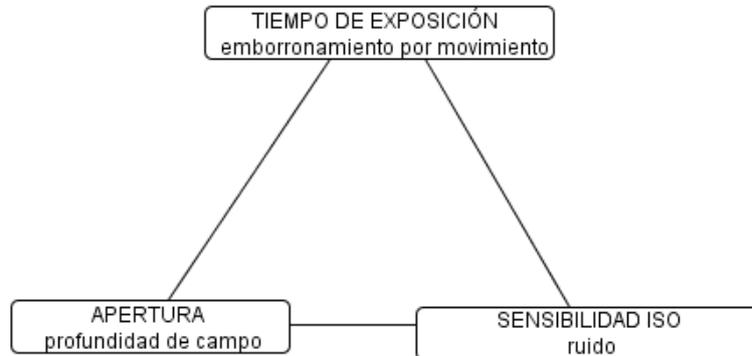


Figura 1.2: Triángulo de exposición. Los tres factores que influyen en la exposición y los efectos que pueden ocasionar.

nativamente para una misma duración bajo la lluvia podríamos utilizar un cubo más estrecho si el objetivo de cantidad de agua recogida fuera menor.

Precisamente los factores que acabamos de enunciar (ancho, tiempo y cantidad) se relacionan con la apertura, tiempo de exposición y sensibilidad ISO que se utilizan en fotografía y que forman el triángulo de exposición de la Figura 1.2. Cada uno de ellos controla la exposición de diferente forma:

- **Apertura:** Controla el área del diafragma por la que dejamos pasar la luz. Se mide en fracciones de número f , la razón entre la distancia focal f y el diámetro de la pupila de entrada. Por ende, un mayor número f indica una apertura de diafragma menor. La reducción de apertura a la mitad del área se conoce como reducción de un paso (f -stop). La mayoría de las cámaras permiten ampliar o reducir el diafragma en saltos de $1/3$ o $1/2$ de paso, lo que significa una mayor precisión en la proporción de luz que atraviesa el diafragma. Un cambio en este parámetro no sólo cambia el área del diafragma que queda abierta sino que modifica la profundidad de campo, o zona que aparece nítida (enfocada), y que coincide con el espacio que permanece en foco delante y detrás del plano enfocado. Una mayor apertura del diafragma (menor número f) reduce la profundidad de campo, provocando que objetos fuera del plano proyecten una imagen borrosa.
- **Tiempo de exposición:** Controla la duración de la exposición; es decir, el tiempo que

permanece abierto el obturador permitiendo que los fotodiodos del sensor acumulen fotones. Consideramos tiempo de exposición corto a valores por debajo de 1/60 segundos. Un tiempo corto produce imágenes oscuras con baja relación señal-ruido (SNR) mientras que una exposición larga produce imágenes más luminosas pero con una alta probabilidad de emborronamiento. Nótese que la velocidad de obturación se define como el inverso del tiempo de exposición [11].

- **Sensibilidad ISO:** También referida como ganancia/factor ISO. La sensibilidad ISO determina cómo de sensible es el sensor de la cámara a la luz. Este concepto viene arrastrado de la fotografía convencional y se conserva en la fotografía digital; aunque sus fundamentos son algo distintos. En fotografía convencional se usaban películas fotográficas formadas por millones de cristales transparentes sensibles a la luz. El tamaño de los cristales es lo que determinaba la sensibilidad de la película. En fotografía digital, es el sensor el encargado de la captación de la imagen. Este sensor lo componen millones de celdas fotosensibles (fotodiodos) formando una malla. La sensibilidad de cada uno de los elementos del sensor es fija, con un valor equivalente a ISO100. Los índices superiores que ofrecen muchas cámaras se logran por una amplificación posterior de la señal captada. Dado que cada una de las celdas genera una cantidad fija de corriente eléctrica aleatoria, aún en ausencia de luz y en relación directa a la temperatura, cuando aumentamos la sensibilidad estamos amplificando la señal lumínica que captan las celdas pero también el ruido. Este ruido se manifestará de forma más evidente en las zonas oscuras de la imagen y en algunos canales más que en otros.

Los mecanismos que hemos enumerado en conjunción con otros tales como el flash, trípode o los sistemas estabilizadores de imagen que incorporan las cámaras DSLR actuales, son útiles para captar imágenes bajo ciertas condiciones. Por desgracia, existen multitud de situaciones en las que la iluminación no es la adecuada y los mecanismos anteriores podrían no ser suficientes para prevenir las degradaciones (véase la Figura 1.3). El flash podría ser insuficiente para fotografiar escenas extensas o a cierta profundidad. Por otro lado, el uso del trípode está condicionado por la superficie de apoyo o la incomodidad de éste. Respecto al estabilizador de la imagen, es una solución útil frente a pequeñas vibraciones de la cámara pero insuficiente si el movimiento es severo. Los problemas para captar imágenes se acentúan si la cámara que utilizamos es una cámara de bajo coste o la cámara de un teléfono móvil ya que estos dispositi-

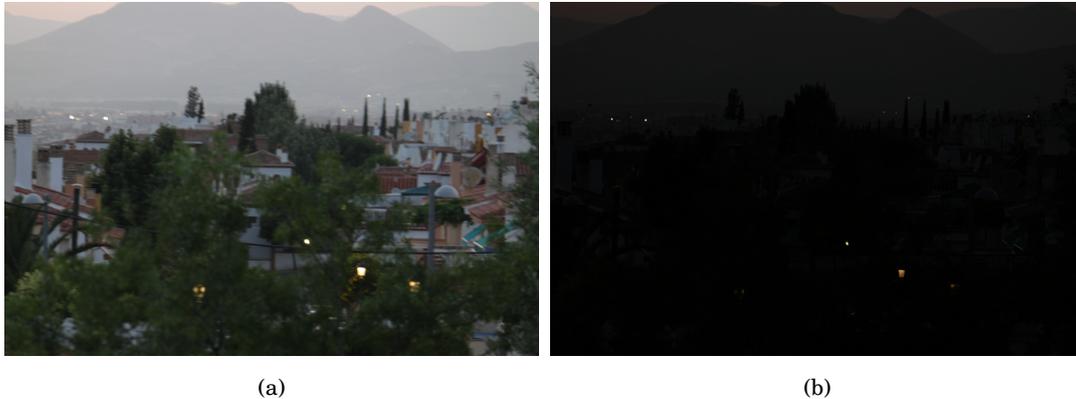


Figura 1.3: Imágenes tomadas en situaciones de baja luminosidad: (a) imagen correctamente expuesta pero borrosa ($f/5.6$, $1/10s$, ISO 100); (b) imagen nítida pero ruidosa ($f/5.6$, $1/160s$, ISO 100).

vos no permiten captar tanta luz como las cámaras SLR, ni suelen proporcionar algunos de los mecanismos descritos anteriormente. En situaciones de escasa luminosidad, es muy probable que ninguno de los mecanismos anteriores sean suficientes para obtener una imagen nítida. Concretamente en esta situación es donde encaja a la perfección el trabajo que presentamos en esta memoria.

El objetivo de esta memoria persigue conseguir una imagen nítida en situaciones de baja luminosidad combinando la información proporcionada por imágenes con distintas exposiciones. En nuestro caso demostraremos que un par de imágenes, formado por una de ellas correctamente expuesta pero borrosa y otra nítida pero muy ruidosa debido al corto tiempo de exposición, como las mostradas en la Fig. 1.3, es suficiente para conseguir una imagen libre de emborronamiento y ruido, mediante la combinación de ambas. Para aprovechar mejor la información de las dos imágenes es necesario conocer el proceso de formación de la imagen que se detalla a continuación.

1.2. Proceso de formación de la imagen

Cuando fotografiamos una escena con una cámara digital para obtener una matriz bidimensional de valores de luminosidad, los valores raramente coinciden con las medidas de radiancia relativas a la escena. Por ejemplo, si un píxel presenta un valor doble que otro, es improbable que se haya observado el doble de radiancia. Es muy frecuente que se aplique una transforma-

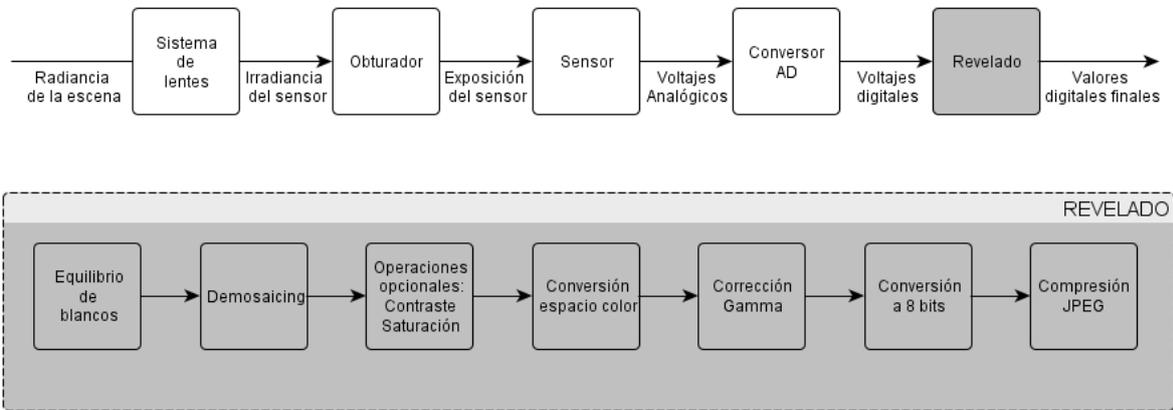


Figura 1.4: Modelo de formación de imagen y revelado digital.

ción no lineal que determina como se convierte la radiancia de la escena en valor de intensidad de imagen. Esta transformación no es fácil de conocer a priori ya que se debe a una concatenación de otras transformaciones no lineales que suceden en el proceso de formación de la imagen. La Figura 1.4 resume las operaciones involucradas en la formación de la imagen final [12]. En este modelo suponemos una fuente emisora de fotones, los cuales viajan siguiendo una línea recta hasta reflejarse en un objeto, describiendo una nueva dirección o son absorbidos por la superficie. La luz que reflejan los objetos (radiancia de la escena) pasa a través del juego de lentes, que dirigen dicha luz hacia el sensor, determinando la irradiancia del sensor, es decir, la cantidad de luz que incide sobre la superficie del sensor medida en W/m^2 . La exposición del sensor se lleva a cabo desde la pulsación del disparador hasta que se agota el tiempo que selecciona el usuario. Durante todo ese tiempo el obturador permanece abierto y el sensor está captando fotones de luz, por tanto la exposición del sensor será igual a la irradiancia del sensor por el tiempo de exposición. El sensor convierte los fotones captados en electrones que son almacenados en fotodiodos hasta que son leídos y las señales analógicas de voltaje sean transformadas en señales digitales usando una precisión usualmente entre 8 y 16 bits. Los sensores actuales suelen proporcionar valores digitales de una profundidad de color de 12 o 14 bits. Una vez obtenidos los valores digitales captados por el sensor, se procede al revelado digital, con el que se obtendrán los valores digitales finales de la imagen. En este paso podemos optar por el revelado directo que realiza la cámara o guardar los datos tal cual los proporciona el sensor (en formato *RAW*) y realizar el revelado digital posteriormente. En cualquiera de los casos, se van a seguir los pasos que aparecen en el cuadro inferior en la Figura 1.4 y que resumimos a

continuación¹:

- **Equilibrio de blancos.** Es la transformación que permite corregir colores dominantes debidos a una iluminación que no es blanca. Esta operación se aplica globalmente multiplicando cada canal por un coeficiente de corrección. Estos coeficientes se conocen como «multiplicadores»; ya que multiplican directamente los valores RGB de cada banda expandiendo o comprimiendo el histograma [11] consiguiendo que los valores de las tres bandas sean iguales en los píxeles que deberían ser grises.
- **Demosaicing.** Habitualmente las cámaras fotográficas captan el color mediante la colocación de una matriz de filtros de color, conocido como máscara Bayer o patrón Bayer, por el que cada 4 píxeles del sensor produce una tripleta de valores RGB. El demosaicing consiste en la interpolación cromática de cada uno de los 4 valores que capta el patrón Bayer colocado delante del sensor de imagen. Después de esta operación obtenemos una tripleta RGB para cada pixel del sensor.
- **Operaciones opcionales.** Es habitual realizar un conjunto de ajustes sobre la imagen para mejorar su apariencia. Debido a la cantidad de operaciones que podrían llevarse a cabo en esta etapa del revelado pasamos a nombrar las más habituales sin detenernos en cada una de ellas.
 - Reducción de ruido
 - Ajustes tonales: Contraste, tonalidad (frío/calor), saturación.
 - Filtro de nitidez o máscara de enfoque.
- **Conversión espacio de color.** La imagen *RAW* se representa en el perfil propio de la cámara, por lo que para una gestión correcta de color del archivo *RAW* es necesario convertirlo a un espacio de color RGB adecuado al dispositivo de representación de la imagen. Si bien hay diferentes espacios de color, es usual utilizar el espacio sRGB (RGB estándar) como espacio de color de salida.

¹Si bien es cierto que otros autores recomiendan un orden diferente en las operaciones de revelado (como en [11] donde se realiza el demosaicing y conversión de espacio de color antes del balance de blancos), nosotros hemos preferido exponer el que realiza el programa DCRAW [13]

- **Corrección Gamma.** Esta operación se encarga de compensar la linealidad de respuesta del sensor corrigiendo la imagen para que se adapte fielmente al sistema de visión humano (SVH), que presenta una respuesta no lineal. Por ejemplo, en comparación con la cámara nuestros ojos son mucho más sensibles a los tonos oscuros que a los tonos de alta luminosidad. Para adaptar la imagen de salida al SVH y representarla correctamente en el dispositivo de salida, aplicamos una corrección gamma que, en el caso de un monitor, será de $1/2,2$. La aplicación de esta corrección gamma se realiza elevando cada uno de los valores RGB lineales, representados como números reales en el rango $[0,1]$, al valor de corrección gamma deseado ($1/2,2$ en nuestro caso). El efecto de esta transformación es el desplazamiento del histograma hacia la derecha, aumentando la luminosidad de la escena.
- **Conversión 8 bits.** El estándar JPEG, al igual que la mayoría de los formatos de almacenamiento de imágenes, define que la precisión para una imagen de color debe ser de 8 bits, por tanto, en caso de trabajar en una precisión mayor tendremos que cuantificar a 8 bits como paso previo a la compresión JPEG.
- **Compresión JPEG.** La última etapa del revelado realiza una compresión con pérdidas, que reduce considerablemente el tamaño de la imagen, conservando el nivel de detalle dependiendo del nivel de compresión aplicado.

Los algoritmos que se detallan en esta memoria trabajarán, normalmente, sobre la imagen que conserva la respuesta lineal del sensor, habitualmente justo detrás del demosaicing, procediéndose al revelado completo una vez terminado el procesamiento.

1.3. Objetivos y estructura de la memoria

Recordamos que el objetivo principal de la memoria va a ser la obtención de una imagen correctamente expuesta y nítida mediante la combinación de un par de imágenes captadas en entornos de baja luminosidad, donde típicamente se obtiene una imagen borrosa pero correctamente expuesta y otra nítida pero subexpuesta, con baja relación señal-ruido. Este objetivo general se desglosa a su vez en el siguiente conjunto de objetivos:

- Calibración geométrica y fotométrica de las imágenes como paso previo a la aplicación del algoritmo de restauración.

- Utilización del paradigma Bayesiano jerárquico para la modelización del problema por su versatilidad y fundamentos teóricos.
- Desarrollo de métodos y técnicas que permitan la combinación de imágenes con distintas exposiciones, minimizando la propagación de emborronamiento y ruido a la restauración.
- Estudiar la combinación de distintos modelos de observación y a priori sobre la imagen.
- Estimación de los parámetros del modelo jerárquico de forma automática.
- Proponer una solución al emborronamiento espacialmente variante.
- Implementar y validar los modelos propuestos, experimentando en imágenes reales y sintéticas, y comparando con otros métodos.

Para la consecución de los objetivos, proponemos la siguiente organización de la memoria: En este primer capítulo hemos presentado el problema de la captación de imágenes en entornos de baja luminosidad, comentado los principales inconvenientes y los mecanismos que incorporan las cámaras para aliviar algunos de los problemas. También se ha expuesto el proceso de formación de imagen y se han fijado los objetivos de la presente memoria.

En el segundo capítulo se hace una revisión de los métodos de restauración a partir de una única imagen y de pares de imágenes. A su vez, se ha profundizado en el tipo de degradaciones que podemos encontrarnos en las imágenes y se propone una taxonomía para organizar las distintas técnicas consideradas. Finalmente se introducen los fundamentos teóricos en los que se apoyan los métodos que vamos a desarrollar: la formulación Bayesiana jerárquica.

El capítulo tercero plantea varias soluciones para realizar la calibración fotométrica y geométrica de las imágenes, como paso previo a la aplicación de cualquiera de los algoritmos de restauración desarrollados.

La parte central de la memoria se compone de los Capítulos 4 y 5. En estos capítulos se proponen los algoritmos de restauración para exposición dual. Más específicamente, en el Capítulo 4 se asume un modelo de emborronamiento espacialmente invariante, mientras que en el Capítulo 5 el modelo propuesto es más realista, destinado a imágenes con emborronamiento espacialmente variante. En todos los algoritmos de restauración propuestos se ha utilizado un paradigma Bayesiano jerárquico; que proporciona un modelo matemático global que nos permite representar nuestro conocimiento acerca de las imágenes e inferir conocimiento sobre

la imagen a estimar. En general, el paradigma Bayesiano jerárquico se compone de una serie de modelos: observación, imagen y emborronamiento, que nos permiten introducir información del conocimiento sobre el problema. Cada uno de ellos depende a su vez, de una serie de parámetros que habrá que estimar junto a la imagen nítida. Sobre dichos parámetros definiremos distribuciones de probabilidad que nos van a permitir la inclusión de información sobre el valor de los parámetros. Este paradigma nos proporciona un marco robusto para incorporar tanto el conocimiento como la incertidumbre de nuestros modelos además de la ventaja de estimar los parámetros asociados de forma automática.

A continuación, en el último capítulo se exponen las conclusiones finales, así como la investigación futura que ha surgido en el desarrollo del trabajo. La memoria concluye con la bibliografía consultada en el trabajo de investigación.

Capítulo 2

Restauración de pares de imágenes: Revisión histórica

Encuentro la televisión muy educativa. Cada vez que alguien la enciende, me retiro a otra habitación y leo un libro.

Julius Henry Marx (Groucho Marx, 1890-1977).
Actor, comediante y escritor, integrante más conocido de los Hermanos Marx.

La restauración de imágenes ha sido un campo muy prolífico desde sus inicios allá por los años 60. El objetivo que se persigue es eliminar la distorsión de la imagen provocada por el emborronamiento y el ruido para recuperar, de forma objetiva, una imagen nítida desconocida. Los dominios de aplicación han sido muy diversos: Exploraciones científicas, investigación legal, vigilancia y seguridad, edición de películas y archivado, Codificación de imágenes y video, fotografía cotidiana.

Uno de los esquemas más utilizado para deshacer la degradación que sufre la imagen es el de primero, modelizar el proceso de degradación, a continuación utilizar información a priori sobre las variables que intervienen en el proceso y aplicar el proceso inverso (deconvolución). Finalmente se podría formular y evaluar de forma objetiva la bondad de la restauración obtenida. Sin embargo, la degradación no es siempre la misma, y el modelo debe ser capaz de adaptarse en cada caso, exigiendo flexibilidad ante distintos tipos de emborronamiento y ruido. Esto hace

el problema aún más atractivo ya que incluso conocida la función de emborronamiento (problema de deconvolución clásica), el proceso de invertir la degradación sufrida es un reto, debido al mal condicionamiento del problema, situación que agrava aún más, si la función de emborronamiento es desconocida (deconvolución ciega), caso que tratamos en esta memoria. El mal condicionamiento del problema considerado se debe a varias razones: primero desconocemos si existe solución; es decir si existe una función de emborronamiento, también conocida como núcleo de emborronamiento o función de dispersión del punto, *Point Spread Function* (PSF), que permita obtener una imagen nítida. Tampoco sabemos si de existir la solución, ésta es única. Es decir, tan válida es la restauración nítida obtenida de una función de emborronamiento que explica perfectamente el movimiento de la cámara como la función de emborronamiento unidad o función delta que produce una restauración borrosa (igual a la observada). Por último, el comportamiento de la solución cambia con una ligera modificación en las condiciones iniciales, es decir en nuestro caso concreto estimaciones levemente diferentes de la función de emborronamiento o variaciones en el ruido de la imagen produce restauraciones totalmente distintas. A continuación mostramos las degradaciones que podrían corromper las imágenes. En primer lugar describimos los tipos de emborronamiento, posteriormente los tipos de ruido que estará siempre presente en nuestras imágenes en mayor o menor medida.

Tipos de emborronamiento

- **Desenfoque:** Se produce frecuentemente en escenas con zonas u objetos a distintas profundidades. Concretamente, la zona que aparece nítida (enfocada), coincide con el espacio en foco delante y detrás del plano enfocado también conocido como profundidad de campo. La profundidad de campo se controla con la apertura del diafragma, de manera que una mayor apertura del diafragma (menor número f) reduce la profundidad de campo. Una apertura grande provoca que objetos fuera del plano proyecten una imagen borrosa.
- **Movimiento de cámara:** Es el emborronamiento más frecuente en imágenes captadas en entornos de baja luminosidad. Se debe fundamentalmente al movimiento involuntario de la cámara por parte del fotógrafo durante la exposición. Nótese que el efecto es más notable cuanto mayor sea la distancia focal del objetivo y por tanto menor el ángulo de visión, lo que, aunque nos permite enfocar objetos a mayor distancia, conduce a una mayor sensibilidad a este tipo de emborronamiento. Algunas cámaras usan técnicas para mitigar

el efecto del emborronamiento por movimiento de cámara tales como el estabilizador de la imagen (OIS-Optical Image Stabilizer del fabricante Canon), cuya función es la de ajustar una de las lentes del conjunto, desplazándola perpendicularmente al eje óptico con el fin de corregir el movimiento de la cámara. Esta solución es útil para pequeñas vibraciones de la cámara pero insuficiente si el movimiento es severo. Otra solución para el emborronamiento por movimiento de cámara, es el uso de un trípode; sin embargo la necesidad de una superficie de apoyo junto a la incomodidad de éste condiciona su uso.

- **Objeto en movimiento:** Es el más complicado de eliminar ya que varía según el movimiento que describa el objeto. La solución del uso de un trípode es ineficaz ante este tipo de emborronamiento; pero si podría aliviarse al considerar un emborronamiento espacialmente variante lo que complica especialmente el proceso de restauración. El Capítulo 5 describe una serie de algoritmos que consideran este tipo de emborronamiento.

La modelización más extendida de la degradación por emborronamiento se realiza en base a la convolución de la imagen nítida \mathbf{x} con una función de emborronamiento, \mathbf{h} , espacialmente invariante, teniendo entonces:

$$\begin{aligned} y_1(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(u-s, v-t)x(s, t)dsdt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(s, t)x(u-s, v-t)dsdt \\ &= h(u, v) * x(u, v), \end{aligned} \quad (2.1)$$

donde $*$ representa el operador de convolución bidimensional. El equivalente discreto a la Ecuación 2.1 para imágenes presenta la siguiente formulación:

$$\begin{aligned} y_1(i, j) &= \sum_{m=1}^M \sum_{n=1}^N h(m, n)x(i-m, j-n) \\ &= \sum_{m=1}^M \sum_{n=1}^N h(i-m, j-n)x(m, n) \end{aligned} \quad (2.2)$$

donde M y N representan el número de filas y columnas de la imagen. Alternativamente podemos expresar la convolución de forma más compacta utilizando la notación matriz-vector, obteniendo:

$$\mathbf{y}_1 = \mathbf{H}\mathbf{x}, \quad (2.3)$$

donde \mathbf{x} e \mathbf{y}_1 son vectores columna de tamaño $MN \times 1$ obtenidos mediante la ordenación lexicográfica de la imagen nítida y borrosa, respectivamente, y \mathbf{H} es una matriz de convolución, de tamaño $MN \times MN$, obtenida a partir de la función de emborronamiento, \mathbf{h} . Si suponemos una convolución circular en la Ec. (2.2), la matriz de emborronamiento \mathbf{H} tiene estructura circulante por bloques. Entonces, la convolución en la parte derecha de la Ec. (2.3) se puede calcular en el dominio de Fourier mediante una multiplicación píxel a píxel de las transformadas de \mathbf{h} y \mathbf{x} .

Cabe destacar que en la realidad se presenta una dificultad añadida, la de encontrarse frecuentemente una combinación de los tres tipos de emborronamiento en una única imagen. Por ello recomendamos tratar el problema desde un marco que contemple el emborronamiento espacialmente variante según la zona.

Tipos de ruido

Además del emborronamiento, el ruido degrada la calidad de la imagen y está siempre presente en dispositivos electrónicos que transmiten y reciben señales. En las cámaras digitales, la señal es la luz que incide sobre el sensor de la cámara. El ruido se debe a una cantidad de factores, aumenta con la sensibilidad ISO que configuremos en la cámara, varía según la duración de la exposición y temperatura e incluso difiere de unos modelos de cámara a otros. A continuación exponemos los tres tipos de ruidos presentes en las cámaras digitales [10].

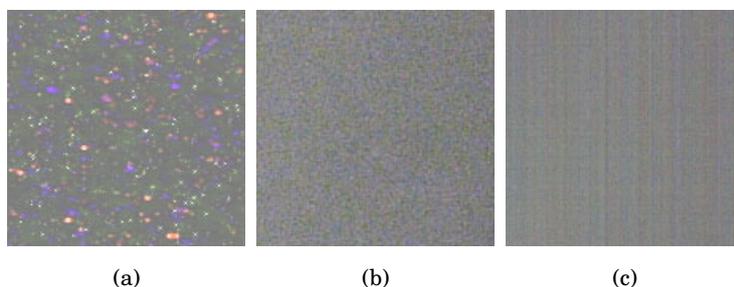


Figura 2.1: Tipos de ruido (de izquierda a derecha): Ruido de patrón fijo, ruido aleatorio y *banding noise*.

Fuente: <http://www.cambridgeincolour.com/>

- **Ruido de patrón fijo:** Incluye tanto píxeles «calientes» (*hot pixels*), como píxeles «muertos» (*dead pixels*). La característica principal de los píxeles calientes es la de presentar un valor saturado ya sea en rojo, verde, azul o combinación de ellos. Por el contrario, los píxeles muertos, están siempre apagados, y de color negro. En ambos casos, la posición de los píxeles defectuosos no varía en diferentes exposiciones o fotografías.
- **Ruido aleatorio:** Se caracteriza por fluctuaciones en color e intensidad de los píxeles sin guardar correlación entre sí (proceso estocástico). Este ruido siempre estará presente en mayor o menor medida, viéndose altamente influenciado por la sensibilidad ISO seleccionada. El patrón de ruido aleatorio cambia incluso cuando seleccionamos los mismos parámetros de exposición.
- **Banding noise:** Es dependiente de la cámara, este tipo de ruido se produce durante la lectura de los datos captados por el sensor. Su presencia se hace más notable en valores altos de sensibilidad ISO y en las sombras, pero también en zonas sobrepuestas. Este tipo de ruido podría empeorarse en ciertos equilibrio de blancos, dependiendo del modelo de cámara.

El ruido de patrón fijo podría suprimirse por ser reproducible. La electrónica de la cámara podría corregirlo conociendo los píxeles afectados por este defecto. Sin embargo, el ruido aleatorio es claramente el más difícil de suprimir sin degradar la imagen. El nivel de ruido también podría cambiar dependiendo de la escena que adquiramos. En las cámaras digitales, las regiones oscuras manifiestan un nivel de ruido más alto que las regiones más luminosas. Estas últimas poseen una señal más robusta debido a una mayor incidencia de la luz, lo que lleva a confirmar que las imágenes de corta exposición tendrán un mayor ruido. Por ende, podría ser ventajoso captar una imagen ligeramente sobrepuesta.

En la comunidad científica se acepta que la degradación producida por el ruido se puede modelar mediante la adición a la imagen ideal que queremos estimar de cierto ruido independiente de la señal. La consideración de ruido aditivo independiente de señal supone una simplificación, ya que algunos tipos de ruidos no están incorrelados con la entrada (ruido de Poisson) e incluso podrían no ser aditivos (ruido multiplicativo). Sin embargo, esta simplificación conduce a métodos de restauración más generales aplicables a un conjunto amplio de problemas. Usando notación matriz-vector, esta degradación puede expresarse como

$$\mathbf{y}_2 = \mathbf{x} + \mathbf{n}_2, \quad (2.4)$$

donde \mathbf{x} denota la imagen nítida desconocida, y \mathbf{n}_2 describe el ruido blanco Gaussiano aditivo independientes de la señal, de tamaño $N \times 1$. El ruido también afecta a las imágenes borrosas por lo que el modelo de imagen de la Ec. 2.3 podría escribirse como

$$\mathbf{y}_1 = \mathbf{H}\mathbf{x} + \mathbf{n}_1, \quad (2.5)$$

donde \mathbf{n}_1 describe el ruido blanco Gaussiano aditivo independientes de la señal, de tamaño $N \times 1$. Las Ecuaciones 2.5 y 2.4 modelan el par de imágenes que usaremos a lo largo de esta memoria, es decir, la imagen correctamente expuesta pero borrosa, \mathbf{y}_1 , y la imagen nítida de corta exposición y, por tanto ruidosa, \mathbf{y}_2 . Nótese que la imagen \mathbf{y}_2 debe estar calibrada fotométrica y geoméricamente con la imagen \mathbf{y}_1 para que los modelos de imagen descritos tengan sentido.

En la siguiente sección revisamos una serie de técnicas que **recuperan una imagen nítida a partir de una única observación, ya sea borrosa o ruidosa**. Seguidamente proponemos una taxonomía de los métodos que combinan pares de imágenes en base al modelado del emborronamiento y se analizan **los métodos más representativos de cada categoría**. Para concluir el capítulo, en la última sección introducimos los fundamentos teóricos que vamos a seguir en el desarrollo de los algoritmos de los Capítulos 4 y 5.

2.1. Métodos de restauración basados en una única observación

En esta sección revisamos algunos de los métodos que restauran la degradación sufrida por una única imagen de forma aislada ya sea emborronamiento o ruido. Para ello analizamos primero las técnicas clásicas y actuales de deconvolución ciega y después los métodos de eliminación de ruido.

Deconvolución ciega

Los modelos clásicos de restauración de imágenes buscan una estimación de la imagen latente asumiendo un núcleo de emborronamiento conocido. En contraposición, la deconvolución ciega trata el problema de forma mucho más realista aunque compleja, ya que la degradación es también desconocida. En general, la degradación es no lineal (incluyendo, saturación y cuantificación) y espacialmente variante (movimiento no uniforme, ópticas imperfectas). Sin embargo para la mayoría de los trabajos, se supone que la imagen observada es la salida de un sistema lineal espacialmente invariante al que se le añade cierto ruido.

En este ámbito existen multitud de técnicas (veáse las revisiones [14, 15]), que difieren principalmente en la información a priori sobre la imagen que se incluye para realizar la restauración. Los primeros algoritmos trataban de identificar la función de emborronamiento a base de patrones [16]. Revisiones más recientes [17, 18, 19] confirman que el área ha sido ampliamente explorada, especialmente en las comunidades de procesamiento de señal, astronómica y óptica. Además, el problema de deconvolución ciega ha sido aplicado en diversidad de escenarios, tales como, imágenes astronómicas [20], imágenes multiespectrales [21], imágenes médicas [22], super-resolución [23], seguimiento [24].

Requiere una mención especial el artículo [25] en el que se analiza y evalúa los distintos enfoques para realizar el proceso de deconvolución ciega. En este trabajo los autores recomiendan la estimación de un emborronamiento suficientemente preciso con el que deconvolucionar, en lugar de estimarlo iterativamente junto a la imagen, ya que esto, amplía la dimensionalidad del problema y lo hace más inestable.

Nuestro interés se centra, principalmente, en los métodos que consideran emborronamiento por movimiento de la cámara. Véase por ejemplo, [26, 27, 28] y las referencias en ellos. En [26], los autores adoptan un marco Bayesiano jerárquico para estimar un núcleo de emborronamiento preciso en un espacio multiescala que evite el mínimo local. La idea general de este enfoque se analiza en la Sección 2.3 de este mismo capítulo. Una vez estimada la función de emborronamiento, los autores utilizan un método de deconvolución no ciega tradicional muy extendido, Richardson-Lucy (RL) [29, 30] para estimar la imagen restaurada. El método RL es una técnica iterativa utilizada originariamente para restauración de imágenes astronómicas contaminadas por ruido de Poisson. Básicamente maximiza la verosimilitud de la restauración usando el algoritmo EM (*Expectation-Maximization*). El algoritmo RL necesita una aproximación fiable del núcleo de emborronamiento. Las bondades de esta técnica son varias: la primera de ellas es la alta velocidad, también la preservación de la energía y la positividad. Como desventaja importante, cabe mencionar la convergencia, normalmente es necesario terminar la ejecución antes de que converja para evitar una amplificación del ruido. Por ello, es frecuente que la condición de parada se base en el ruido residual. El algoritmo propuesto en [26], necesita intervención del usuario para seleccionar un bloque con alto contenido en fronteras para inicializar la PSF. También es necesario inicializar la dirección de la función de emborronamiento, ya sea en horizontal o vertical, el tamaño de ésta y los parámetros de las distribuciones a priori. En [27] los autores pretenden aliviar el efecto de anillado que padecen las restauraciones obtenidas de una única

imagen borrosa. Por ello, proponen un modelo de ruido más efectivo junto a una mejor gestión de los artificios visuales causados por errores en la estimación del núcleo de emborronamiento. Esto último se lleva a cabo mediante la imposición de una restricción de suavizado sobre imagen latente en áreas de contraste bajo en la observación borrosa. La tercera de las contribuciones se centra en la optimización del algoritmo mediante una serie de mejoras. El método en [28], presenta un algoritmo que estima las incógnitas (imagen nítida, emborronamiento e hiperparámetros) de manera automática y simultánea empleando un marco Bayesiano jerárquico. Para ello, se apoya en una serie de modelos que aproximan el conocimiento que tenemos de las incógnitas, así para el conocimiento a priori de la imagen se utiliza el modelo TV, y SAR como modelo de emborronamiento.

Recientemente están apareciendo métodos capaces de modelar un emborronamiento espacialmente variante a partir de una única observación borrosa. En [31], los autores extienden el modelo en [32] para eliminar emborronamiento espacialmente variante a una única imagen borrosa. El algoritmo que proponen se compone de cuatro fases. En la primera etapa (preprocesamiento) se realzan los bordes, suprime ruido y calcula los gradientes más significativos. Posteriormente en una segunda fase, se calcula el núcleo de emborronamiento. La imagen nítida se estima utilizando este núcleo de emborronamiento estimado. Finalmente, los autores proponen un proceso de refinamiento en el que se estima la imagen nítida final, mediante la adaptación del método en [33]. Debido a la dificultad del problema, el algoritmo presenta una serie de limitaciones en caso de objetos en movimiento, desenfoque, píxeles saturados y ruido. Además, como en la mayoría de métodos de deconvolución ciega que utilizan una única observación, este método no es aplicable a imágenes degradadas por emborronamiento severo y necesita de la supervisión del usuario en las fases iniciales para facilitar la convergencia del algoritmo. Por ende, los métodos que consideran una segunda observación permiten aliviar todos los inconvenientes anteriores y serán revisados en la siguiente sección.

Eliminación de ruido

Una forma de aliviar el posible emborronamiento en una escena de baja luminosidad consiste en seleccionar un tiempo de exposición menor; sin embargo esto lleva asociado un incremento considerable del nivel de ruido más acusado aún en escenas de baja luminosidad. Algunos de los tipos de ruido frecuentes que podemos encontrarnos en las fotografías que realizamos de forma cotidiana ya han sido descritos en la Sección 2 del presente capítulo. El reto al que se enfrentan

las técnicas de eliminación de ruido consiste en reducir el nivel de ruido pero produciendo una imagen nítida preservando los detalles. Existen abundantes métodos para supresión de ruido y aún más variantes que tratan de mejorarlos. En [34] se realiza una revisión muy completa de los algoritmos de eliminación de ruido. Además, los autores proponen un nuevo algoritmo NL-means (Non Local means) que consigue preservar la estructura de una imagen digital y lo evalúan frente a las mejores versiones de los métodos de eliminación de ruido. El algoritmo NL-means explota la similitud entre bloques cercanos a uno dado que pueden verse menos afectados por el ruido, y por tanto, sería apropiado sustituir bloques ruidosos con otros donde el ruido es menor. Los resultados son excepcionales en imágenes periódicas (que presentan un patrón que se repite). Aparte de NL-means y los algoritmos comparados en [35] también merecen especial interés el método en [36] basado en *wavelet*, que fue considerado el estado del arte hasta hace poco, donde los autores aplican la eliminación de ruido mediante el modelado de los coeficientes *wavelet* de la imagen como una mezcla de Gaussianas recibiendo como uno de los parámetros del algoritmo una aproximación de la varianza del ruido. Otro método más actual (BM3D) [1], considerado el estado del arte en eliminación de ruido. BM3D es una aproximación novedosa al filtrado no paramétrico adaptativo no local. El método propuesto en [1] adopta una aproximación híbrida la cual agrupa bloques similares y aplica un filtrado colaborativo en el dominio de la transformada de coseno discreta (DCT *discrete cosine transform*). El dominio transformado permite distinguir la señal del ruido más fácilmente que en el dominio espacial.

2.2. Métodos basados en la combinación de observaciones

En general, las técnicas que se apoyan en una segunda observación de menor exposición, necesitan un menor ajuste de los parámetros iniciales del algoritmo. Además permiten eliminar emborronamientos más complejos que utilizando una única imagen. Desafortunadamente, también añaden una serie de dificultades, tales como la desalineación entre las imágenes y el ruido de la observación de corta exposición. En esta sección, enumeramos algunos de los métodos que estiman la imagen nítida a partir de la combinación de una observación borrosa y una segunda observación nítida pero ruidosa en exceso, es decir, lo que denominamos exposición dual o par de imágenes. Para ello proponemos la siguiente taxonomía: primeramente analizamos las técnicas que asumen un único núcleo de emborronamiento para toda la imagen, es decir, el emborronamiento es considerado espacialmente invariante y por tanto son métodos menos

flexibles. Seguidamente analizamos los métodos que asumen la función de emborronamiento dependiente de la posición, esta degradación es más próxima a la realidad y se conoce como emborronamiento espacialmente variante. Concluiremos la sección con algunos algoritmos de fusión de imágenes, que no precisan estimar la PSF.

2.2.1. Métodos basados en PSF espacialmente invariante

Los métodos que analizamos en este apartado [37, 38, 2, 5], se caracterizan por modelizar el proceso de emborronamiento como un único núcleo global para toda la imagen, el cual es estimado y utilizado para invertir el proceso de degradación, con la intención de obtener una imagen sin emborronamiento ni ruido.

El método en [37] propone realizar la calibración de un par de imágenes basándose en las propiedades del emborronamiento estimado en distintas escalas. Los autores aprovechan que la alineación de las imágenes influye directamente en la PSF, para encontrar la mejor alineación que produce el núcleo de emborronamiento de mayor rareza (en términos de mayor número de ceros). El núcleo de emborronamiento estimado podría ser utilizado para deconvolucionar la imagen borrosa. Sin embargo el método de registrado propuesto no es robusto a transformación de perspectiva o movimientos no globales. En [38], la imagen de exposición normal y la versión suavizada de la imagen de corta exposición se utilizan para estimar la PSF. Sin embargo, es necesario un conocimiento preciso de la varianza del ruido en la imagen de corta exposición para aplicar el método de eliminación de ruido, ya que una varianza de ruido incorrecta, podría provocar un suavizado en exceso y por consiguiente supresión de detalles que se propagarían a la restauración final. En [2, 5], se propone un marco Bayesiano para modelar la imagen desconocida y la función de emborronamiento, así como los parámetros asociados a los modelos. El uso de este marco, permite estimar la imagen, el emborronamiento y los parámetros sin intervención del usuario. En [5], describimos un procedimiento para combinar tres modelos de observación correspondientes a la imagen de exposición normal, la imagen de corta exposición y una combinación del modelo de las imágenes con distintas exposiciones. Los detalles de este método serán expuestos en el Capítulo 4.

2.2.2. Métodos basados en PSF espacialmente variante

Los trabajos comentados en el apartado anterior, suponen un emborronamiento invariante en toda la imagen, lo cual ocurre con poca frecuencia en la práctica. Se ha demostrado recien-

temente que incluso el sistema óptico de la cámara genera un considerable emborronamiento espacialmente variante, más acusado en los bordes de la imagen [39]. Por tanto la suposición previa de emborronamiento invariante en la imagen podría dar lugar a artificios importantes en algunas regiones o suavizados excesivos en otras.

En esta sección revisamos una serie de trabajos recientes, que permiten superar las limitaciones del emborronamiento espacialmente invariante, mediante la introducción de un modelo de emborronamiento espacialmente variante que se caracteriza por la modelización de la variación del emborronamiento a lo largo de la imagen. Así, en [32], los autores desarrollan un modelo de emborronamiento paramétrico no uniforme a partir de un único núcleo que permite modelar una gama más amplia de emborronamientos aplicando transformaciones espaciales (globales) sobre el núcleo de partida. Este enfoque permite modelar un emborronamiento espacialmente variante aunque su aplicación es exclusiva para movimiento de la cámara no permitiendo corregir, al mismo tiempo, otras formas de emborronamiento como desenfoque. El modelo que proponen se aplica sobre un par de imágenes borrosa y ruidosa pero también proponen otra versión del algoritmo para aplicarlo sobre una única imagen borrosa.

El uso de una imagen con baja exposición adicional ayuda a solventar los problemas descritos anteriormente. Así, en [40, 4] el emborronamiento espacialmente variante se modela y estima de la siguiente forma: inicialmente las imágenes se dividen en bloques independientes y se estima un núcleo de emborronamiento en cada uno de ellos. Posteriormente se aplica el proceso de deconvolución haciendo uso del núcleo estimado en cada bloque. El proceso de deconvolución propuesto en [40] fue posteriormente mejorado en [4] mediante la inclusión de un término en el proceso de restauración que incluye información de la imagen de corta exposición. Las restauraciones obtenidas usando los métodos en [40, 4] pueden padecer un indeseado efecto de bloques en parches adyacentes debido a cambios en el emborronamiento. Además, algunos parámetros de los algoritmos deben ser seleccionados por el usuario manualmente, principalmente el umbral para detección de núcleos erróneos.

Los métodos en [6, 7], que forman parte de la aportación de esta tesis y desarrollamos ampliamente en el Capítulo 5, también usan el procesamiento en bloques independientes y persiguen aliviar las limitaciones de los métodos anteriores incorporando un novedoso mecanismo que selecciona, para cada bloque en que dividimos las imágenes de entrada, entre la restauración obtenida por deconvolución a partir del par de imágenes y la restauración por eliminación de ruido de la imagen de corta exposición. Nótese que en el primero de los casos se utilizó la

información de ambas observaciones mientras que en el caso de restauración por eliminación de ruido únicamente se hace uso de la imagen de corta exposición y se aplicará en bloques en los que, la función de emborronamiento no se estimó con suficiente precisión y por tanto es preferible desechar la información del bloque borroso. El proceso de deconvolución propuesto hace uso de un modelo Bayesiano jerárquico en el que se estiman de forma automática todos los parámetros de los modelos. Además la restauración final no sufre el efecto de bloques por el empleo de bloques solapados junto a una función de enventanado para la fusión de las diferentes estimaciones para cada bloque.

Métodos de fusión sin estimación de PSF

En esta sección se describen métodos de fusión de imágenes [41, 3, 42, 43] que evitan el cálculo explícito de la función de emborronamiento y por tanto prescinden de la etapa de deconvolución. En [41] los autores aprovechan la información del color y las restricciones espaciales simultáneamente. Adoptan un marco Bayesiano para incorporar esta información y generar una función óptima de mapeo de color la cual hace corresponder los colores de la imagen de corta exposición a los de la imagen de exposición normal. Esta técnica permite su aplicación a una variedad de emborronamiento pero depende en exceso de la calidad de la imagen de baja exposición, concretamente en caso de ruido severo éste será propagado a la imagen resultado. El método en [3] clasifica los píxeles de las dos imágenes como borrosos o nítidos. A continuación éstos son fusionados mediante una combinación lineal de ambos. En un trabajo más reciente [42], los autores proponen un algoritmo aplicable a un conjunto de imágenes con diferentes tiempos de exposición captadas por una cámara de móvil. El algoritmo tiene como objetivo la obtención de una imagen (probablemente borrosa) de alto rango dinámico (HDR) que fusionan con una imagen nítida pero ruidosa debido a la corta exposición. Para la fusión utilizan una descomposición wavelet seleccionando, en cada escala y orientación, la información de la imagen HDR que va a ser incluida en la ruidosa con el fin de obtener una imagen HDR libre de emborronamiento. En [43], se propone un algoritmo de fusión de imágenes basado en métricas de la importancia de regiones que fusiona, de acuerdo con algunas características tales como el contraste, el tamaño y forma de la región en cada una de las imágenes.

2.3. Formulación Bayesiana jerárquica

El paradigma bayesiano jerárquico ha venido empleándose en muchas áreas relacionadas con el análisis de imágenes. Algunos ejemplos son: *compressive sensing* [44], super-resolución [45], clasificación [46], medicina [22], restauración [47, 48].

La modelización bayesiana de un problema se basa en una aproximación estadística y está relacionada con la teoría de la decisión en presencia de conocimiento estadístico que pueda arrojar luz sobre la incertidumbre involucrada en los problemas de decisión. La estadística clásica se dirige hacia el uso de la información proveniente de los datos obtenidos de investigación estadística para hacer inferencias sobre datos desconocidos. La teoría de la decisión, por otra parte, intenta combinar la información de los datos con otros aspectos relevantes del problema para tomar decisiones mejores. Otro punto de vista del mismo problema, desde la teoría de la regularización, se puede estudiar en [49, 50]. Esta teoría persigue penalizar las soluciones que se alejen de la deseada mediante la inclusión de un término al que llamamos regularizador.

El enfoque conocido como Inferencia Bayesiana se puede distinguir de otros enfoques estadísticos por el uso de información a priori sobre los modelos planteados. Aplicado a restauración de imágenes, esta información a priori suele contener información espacial sobre la distribución de las intensidades de los objetos de la imagen. Este enfoque ha tenido una amplia aplicación en el campo de la restauración de imágenes en Astronomía (véase, por ejemplo, [20], y las referencias allí citadas), donde las imágenes se consideran compuestas de fuentes puntuales (estrellas) y objetos cuya luminosidad varía suavemente (planetas y galaxias), aunque ha demostrado su efectividad en otras aplicaciones de la restauración de imágenes (véase, por ejemplo, [47, 48]).

2.3.1. Descripción del paradigma bayesiano para exposición dual

En el paradigma bayesiano, distinguimos entre \mathbf{x} , la imagen nítida desconocida, que se habría observado bajo condiciones de captación ideales, la función de emborronamiento \mathbf{h} , e \mathbf{y} , un par de observaciones de la misma escena con distintas exposiciones. Estas observaciones ya se describieron en las Ecs. 2.5 y 2.4.

El objetivo es reconstruir \mathbf{x} , la imagen original, junto a la función de emborronamiento \mathbf{h} , a partir de las observaciones \mathbf{y} . Los métodos bayesianos parten de una distribución a priori, una distribución de probabilidad sobre la imagen, $p(\mathbf{x}|\alpha)$. En esta distribución a priori es donde se incorpora información sobre la estructura esperada en una imagen. Equivalentemente

se define una distribución a priori para la PSF $p(\mathbf{h}|\gamma)$. También resulta necesario especificar $p(\mathbf{y}|\mathbf{x}, \mathbf{h}, \beta)$, la distribución de probabilidad condicionada de las imágenes observadas, \mathbf{y} , si \mathbf{x} fuera la verdadera imagen y \mathbf{h} el emborronamiento real. Tanto las distribuciones a priori como la condicionada dependen de una serie de parámetros, conocidos como hiperparámetros, y que denotamos por α , β y γ . Estos hiperparámetros habrá que fijarlos o estimarlos a partir de los datos. El paradigma jerárquico bayesiano combina la información proveniente de los datos de la imagen observada por el sistema de captación con la información a priori para construir la distribución a posteriori de \mathbf{x} dada \mathbf{y} , el conjunto de observaciones y \mathbf{h} , la PSF, $p(\mathbf{x}|\mathbf{y}, \mathbf{h}, \alpha, \beta, \gamma)$, a partir de la cual se toman decisiones y se desarrollan inferencias.

El paradigma jerárquico bayesiano aplicado al ámbito de la restauración de imágenes digitales tiene, al menos, dos fases. En la primera, se emplea el conocimiento sobre las características del ruido, el emborronamiento y el comportamiento estructural de la imagen estimada para formar $p(\mathbf{y}|\mathbf{x}, \mathbf{h}, \beta)$, $p(\mathbf{h}|\gamma)$ y $p(\mathbf{x}|\alpha)$, respectivamente. En la segunda fase, el paradigma jerárquico bayesiano establece una distribución a priori sobre los hiperparámetros desconocidos, $p(\alpha)$, $p(\beta)$ y $p(\gamma)$, donde se incluye nuestro conocimiento subjetivo sobre el posible valor de los hiperparámetros. Aunque, a veces, es posible conocer relaciones entre los hiperparámetros, nosotros estudiaremos aquí el modelo donde la distribución conjunta de probabilidad queda definida como [50]

$$p(\mathbf{y}, \mathbf{x}, \mathbf{h}, \Theta) = p(\alpha)p(\beta)p(\gamma)p(\mathbf{x}|\alpha)p(\mathbf{h}|\gamma)p(\mathbf{y}|\mathbf{x}, \mathbf{h}, \beta),$$

con $\Theta = \{\alpha, \beta, \gamma\}$.

2.3.2. Soluciones al problema de la restauración

Una vez definidos los elementos necesarios para realizar el análisis basado en el paradigma jerárquico bayesiano, podemos realizar este análisis de varias forma: por ejemplo, usando el análisis basado en la evidencia, el basado en la moda a posteriori (MAP), o la aproximación variacional al análisis Bayesiano. A continuación mostramos los tres enfoques:

Análisis basado en la evidencia

En el análisis basado en la evidencia [50] primero se integra la distribución conjunta respecto de \mathbf{x} y \mathbf{h} , para estimación de los hiperparámetros. Posteriormente utilizando estas estimaciones se obtiene la estimación de la PSF y la imagen empleando los hiperparámetros estimados. El proceso de estimación se resume a continuación:

- Estimación de los hiperparámetros:

$$\hat{\Theta} = \arg \max_{\Theta} p(\Theta | \mathbf{y}) = \arg \max_{\Theta} \int_{\mathbf{x}, \mathbf{h}} p(\mathbf{y}, \mathbf{x}, \mathbf{h}, \Theta) d\mathbf{x} d\mathbf{h}.$$

- Estimación de la función de emborronamiento e imagen:

$$\hat{\mathbf{x}}, \hat{\mathbf{h}} = \arg \max_{\mathbf{x}, \mathbf{h}} p(\mathbf{x} | \mathbf{y}, \mathbf{h}, \hat{\Theta}).$$

Análisis basado en la moda a posteriori

Este enfoque se basa en la maximización de la distribución a posteriori $p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})$ o estimador MAP (*Maximum a Posteriori*), y realiza la estimación de la imagen, el emborronamiento y los hiperparámetros de forma simultánea integrando $p(\mathbf{x}, \mathbf{y}, \mathbf{h}, \Theta)$ sobre los hiperparámetros Θ para obtener la verosimilitud y maximizarla sobre \mathbf{x} e \mathbf{y} . El proceso de estimación de la imagen, \mathbf{x} , el emborronamiento \mathbf{h} y de los hiperparámetros, Θ , se describe a continuación:

- Estimación de la función de emborronamiento y de la imagen:

$$\hat{\mathbf{x}}, \hat{\mathbf{h}} = \arg \max_{\mathbf{x}, \mathbf{h}} p(\mathbf{x}, \mathbf{h}, \mathbf{y}) = \arg \max_{\mathbf{x}} \int_{\Theta} p(\mathbf{y}, \mathbf{x}, \mathbf{h}, \Theta) d\Theta.$$

- Estimación de los hiperparámetros:

$$\hat{\Theta} = \arg \max_{\Theta} p(\Theta | \mathbf{y}, \hat{\mathbf{h}}, \hat{\mathbf{x}}).$$

Aproximación variacional al análisis Bayesiano

El procedimiento de inferencia persigue optimizar una función dada, y no en obtener la distribución a posteriori, que puede ser analizada o simulada para obtener información adicional sobre la bondad de la estimación. En lugar de considerar una distribución sobre todos los posibles valores de los parámetros junto con la imagen y el emborronamiento, los procesos de estimación antes descritos seleccionan un conjunto de valores específicos. Esto significa que se han obviado otras muchas de las interpretaciones de los datos. Si la distribución a posteriori presenta un pico pronunciado, otros valores para los hiperparámetros, la imagen y el emborronamiento tendrán una probabilidad mucho menor; pero si la distribución es más ancha, la elección de un único valor va a descartar otras muchas soluciones con una probabilidad muy parecida.

Los métodos variacionales proporcionan una forma de aproximar la distribución a posteriori $p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})$ por una distribución más sencilla, de la que es más fácil extraer observaciones. Véanse las tesis [51, 52], libros [53, 54] y el capítulo del libro [55] para una introducción más completa de los métodos variacionales.

En los últimos años ha crecido el interés en la aplicación de métodos variacionales a distintos problemas de inferencia. Estos métodos tratan de aproximar las distribuciones a posteriori utilizando la divergencia de Kullback-Leibler [56, 57]. Esta teoría se ha aplicado a un conjunto de problemas muy amplio: modelos gráficos y redes neuronales [55], análisis de componentes principales (ICA) [51], modelos de Markov [52], SVM [58], super-resolución [59] y deconvolución ciega [60]. Como ya hemos comentado previamente, la inferencia en $\mathbf{x}, \mathbf{h}, \Theta$ debe basarse en la distribución a posteriori $p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})$. Pero hay ocasiones en que ésta no puede calcularse de forma explícita, pero podemos aplicar los métodos variacionales para aproximar la distribución por otra distribución $q(\mathbf{x}, \mathbf{h}, \Theta)$ que es una simplificación de la distribución a posteriori real y que nos permite realizar la marginalización necesaria en el proceso de inferencia.

El proceso utilizado para encontrar la aproximación $q(\mathbf{x}, \mathbf{h}, \Theta)$ consiste en la minimización de la divergencia de Kullback-Leibler, dada por,

$$\begin{aligned} C_{KL}(q(\mathbf{x}, \mathbf{h}, \Theta) \| p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})) &= \int q(\mathbf{x}, \mathbf{h}, \Theta) \log \left(\frac{q(\mathbf{x}, \mathbf{h}, \Theta)}{p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})} \right) d\mathbf{x} d\mathbf{h} d\Theta \\ &= \int q(\mathbf{x}, \mathbf{h}, \Theta) \log \left(\frac{q(\mathbf{x}, \mathbf{h}, \Theta)}{p(\mathbf{y}, \mathbf{x}, \mathbf{h}, \Theta)} \right) d\mathbf{x} d\mathbf{h} d\Theta + \text{const}, \end{aligned} \quad (2.6)$$

que es siempre positiva e igual a cero únicamente si $q(\mathbf{x}, \mathbf{h}, \Theta) = p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})$. Para abordar este cálculo es muy frecuente asumir que la distribución $q(\mathbf{x}, \mathbf{h}, \Theta)$ es factorizable [54], es decir:

$$q(\mathbf{x}, \mathbf{h}, \Theta) = q(\mathbf{x})q(\mathbf{h})q(\Theta), \quad (2.7)$$

considerando que las distribuciones sobre la imagen $q(\mathbf{x})$, el emborronamiento $q(\mathbf{h})$ y los parámetros desconocidos $q(\Theta)$ son independientes. Ahora procedemos a encontrar la mejor de las distribuciones en términos de divergencia. Usando la Ec. 2.7 en la Ec. 2.6, obtenemos la siguiente divergencia

$$\begin{aligned} C_{KL}(q(\mathbf{x}, \mathbf{h}, \Theta) \| p(\mathbf{x}, \mathbf{h}, \Theta | \mathbf{y})) &= \int_{\Theta} q(\Theta) \left(\int_{\mathbf{x}, \mathbf{h}} q(\mathbf{x})q(\mathbf{h}) \log \left(\frac{q(\mathbf{x}, \mathbf{h}, \Theta)}{p(\mathbf{y}, \mathbf{x}, \mathbf{h}, \Theta)} \right) d\mathbf{x} d\mathbf{h} \right) d\Theta + \text{const}, \\ &= \int_{\mathbf{x}, \mathbf{h}} q(\mathbf{x})q(\mathbf{h}) \left(\int_{\Theta} q(\Theta) \log \left(\frac{q(\mathbf{x}, \mathbf{h}, \Theta)}{p(\mathbf{y}, \mathbf{x}, \mathbf{h}, \Theta)} \right) d\Theta \right) d\mathbf{x} d\mathbf{h} + \text{const}, \end{aligned} \quad (2.8)$$

Ahora dada $\hat{q}(\Theta)$ una estimación de $q(\Theta)$, podemos obtener como estimación de $q(\mathbf{x})$ y $q(\mathbf{h})$ el resultado de

$$\hat{q}(\mathbf{x}), \hat{q}(\mathbf{h}) = \arg \min_{q(\mathbf{x})q(\mathbf{h})} C_{KL}(\hat{q}(\Theta)q(\mathbf{x})q(\mathbf{h}) \parallel p(\mathbf{x}, \mathbf{h}, \Theta|\mathbf{y})), \quad (2.9)$$

y dada $\hat{q}(\mathbf{x})$ y $\hat{q}(\mathbf{h})$, la estimación de $q(\Theta)$ y $q(\mathbf{h})$, podemos obtener una estimación de $q(\Theta)$ mediante

$$\hat{q}(\Theta) = \arg \min_{q(\Theta)} C_{KL}(q(\Theta)\hat{q}(\mathbf{x})\hat{q}(\mathbf{h})) \parallel p(\mathbf{x}, \mathbf{h}, \Theta|\mathbf{y}), \quad (2.10)$$

Capítulo 3

Calibración de las imágenes

Si buscas resultados distintos, no hagas siempre lo mismo.

Albert Einstein (1879-1955).

Físico. Considerado el científico más importante del siglo XX.

En este capítulo se analizan varios métodos para la calibración automática del par de imágenes con distintas exposiciones captadas de una misma escena en tiempos ligeramente diferentes. Cuando las imágenes se realizan con distintos parámetros de exposición se produce un desajuste en la luminosidad y además al captarlas en instantes distintos muy probablemente no se encuentren alineadas, es decir la información para la misma posición no coincide en ambas imágenes. Con el fin de aplicar cualquiera de los algoritmos de restauración es necesario corregir estas distorsiones primero. Es por ello, que la etapa de calibrado va a servir como preprocesamiento para cualquiera de los algoritmos de restauración que deseemos aplicar y que comentamos en detalle en los Capítulos 4 y 5. La etapa de calibrado fotométrico tiene como finalidad ajustar la luminosidad de las imágenes, mientras que el objetivo del calibrado geométrico es el de registrar las imágenes, y por tanto conseguir que coincidan las posiciones de las imágenes. Existen multitud de métodos en la literatura para el calibrado de imágenes [61, 62, 63, 64]. En esta memoria vamos a exponer los que hemos considerado más efectivos para nuestro problema concreto.

El primer paso para el calibrado va a consistir en la transformación de espacio de color de las imágenes de RGB a YC_bC_r , donde Y representa la banda de luminancia y las otras dos

bandas representan la diferencia entre bandas de color azul y roja con la luminancia Y . Debemos hacer notar que los algoritmos de restauración trabajarán sobre la banda de luminancia y, por tanto ésta es la banda que debemos calibrar. La restauración final va a estar formada de la luminancia restaurada y las bandas de color de la imagen borrosa. La justificación de restaurar la banda de luminancia se debe a que los colores de la imagen oscura se han perdido o están totalmente distorsionados por lo que no aportarán información para la restauración de la imagen. Por ello en este capítulo vamos a centrarnos en la calibración de la banda de luminancia. El procedimiento general sería aplicar un calibrado fotométrico de las bandas de luminancia y posteriormente calibrar geoméricamente las imágenes. Sin embargo debemos puntualizar que existe una gran dependencia entre el calibrado fotométrico y geométrico, ya que en algunos casos será necesario aplicar uno antes que otro o viceversa, o incluso aplicar un segundo calibrado fotométrico más fino tras calibrar geoméricamente las imágenes de entrada originales. Las ventajas de aplicar un segundo calibrado fotométrico son, en general, muy pequeñas y en concreto, para la técnica de ecualización de histograma, que es la técnica que propondremos usar para el calibrado fotométrico, prácticamente nulas ya que esta técnica es independiente a la geometría de la escena. Sin embargo, al ajustarse el histograma de las mismas porciones de la imagen, después del registrado, podría producir un ajuste de luminosidad ligeramente más fino.

3.1. Calibración fotométrica

El calibrado fotométrico, consiste en equilibrar la luminosidad de la imagen de corta exposición para ajustarla al mismo nivel que la imagen de exposición correcta (aunque borrosa). Con este objetivo podemos utilizar cualquiera de las siguientes técnicas

- Ajuste de la ganancia. (Sec. 3.1.1)
 - por método directo.
 - por mínimos cuadrados.
- Ecualización de histograma mediante histograma de imagen borrosa. (Sec. 3.1.2)

El método elegido para la calibración fotométrica ha sido la ecualización de histograma, que como veremos más adelante proporciona un mejor ajuste de iluminación que el resto de técnicas

y posee dos ventajas claras: la primera de ellas se debe a que esta técnica es independiente de la geometría de la escena, no necesitando una calibración geométrica a priori, y la segunda es que podemos introducir fácilmente la información de la imagen borrosa para ayudarnos a calibrar la imagen oscura.

3.1.1. Ajuste de la ganancia

Los dos métodos que explicamos en este apartado van a calcular un factor de ganancia (λ) que deberá ser aplicado sobre la imagen oscura para incrementar la luminosidad. El primer inconveniente que nos encontramos en este tipo de técnicas es que se realiza un calibrado fotométrico global; es decir el parámetro estimado es el mismo para todos los píxeles de la imagen, por tanto va a ser necesario aplicarlo sobre un entorno que sea lineal, para que aumenten en la misma proporción. Esto se consigue deshaciendo la curva de respuesta de la cámara permitiéndonos operar en el dominio lineal del sensor. Para ello, seguimos el mismo procedimiento que en [65], donde los autores sugieren utilizar un valor de $\gamma = 2$ cuando la curva de respuesta es desconocida. Esta curva coincide con la corrección Gamma que describimos en el capítulo primero y que debemos deshacer. Nótese que la corrección se aplica siempre en el dominio $[0, 1]$, así que previamente deberemos dividir por el máximo valor que podamos tener en las imágenes, en nuestro caso por ser imágenes de 8 bits será 255. Para volver al dominio no lineal basta con aplicar la curva inversa (γ^{-1}) y posteriormente nos llevaremos las imágenes al rango correcto multiplicando por 255. Nótese que, de trabajar con imágenes *RAW*, es posible revelar la imagen sin realizar la corrección gamma y aplicar dicha corrección una vez realizado el proceso de restauración.

Método directo

En este método, la ganancia λ se calcula directamente de los parámetros de la cámara (ISO y tiempo de exposición) que se usaron para captar cada una de las imágenes. Siguiendo el procedimiento en [65] y asumiendo que la apertura del diafragma se mantiene constante, la ganancia se obtiene como

$$\lambda = \frac{ISO_{y_1} \cdot t_{y_1}}{ISO_{y_2} \cdot t_{y_2}}, \quad (3.1)$$

donde ISO_{y_1} , ISO_{y_2} , t_{y_1} y t_{y_2} representan el valor ISO y tiempo de exposición de cada una de las imágenes borrosa y ruidosa, respectivamente. Recordemos que el valor ISO controla la

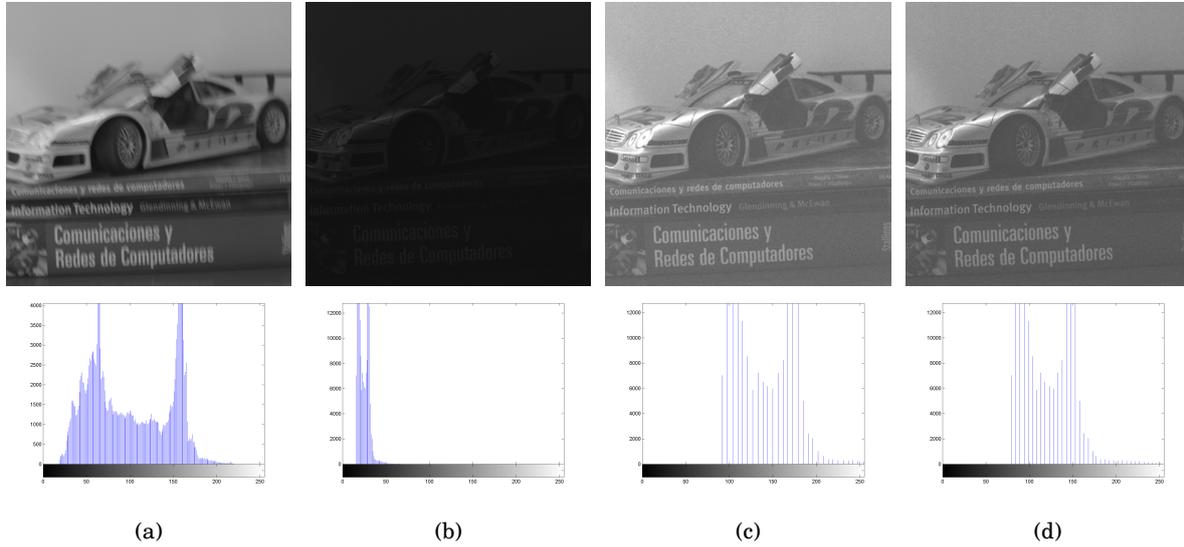


Figura 3.1: (a) Imagen borrosa e histograma. (b) Imagen de corta exposición con su histograma. (c) Resultado de aplicar método directo estimación de ganancia ($\lambda = 33,3$), debajo su histograma. (d) Resultado del calibrado fotométrico de aplicar método de mínimos cuadrados para la estimación de ganancia ($\lambda = 24,3$), debajo su histograma.

ganancia o amplificación que se va a aplicar sobre cada una de las celdas del sensor y el tiempo de exposición indica la fracción de tiempo que permanece el obturador abierto.

Ajuste por mínimos cuadrados

Del mismo modo que el método anterior, se va a estimar un factor de ganancia que aplicado sobre la imagen de corta exposición, en el dominio de irradiancia, nos permita incrementar la luminosidad global de la imagen. Sin embargo, en esta ocasión el parámetro de ganancia se estima mediante mínimos cuadrados asumiendo que las imágenes se diferencian por un único factor de ganancia (λ). Así, λ se calcula como

$$\lambda = \arg \min_{\lambda} \sum_{n=1}^N [y_1(n) - \lambda y_2(n)]^2, \quad (3.2)$$

donde N expresa el tamaño de la imagen, resultando en un estimador de λ

$$\lambda = \frac{\sum_{n=1}^N [y_1(n)y_2(n)]}{\sum_{n=1}^N y_2(n)^2}. \quad (3.3)$$

Este proceso presenta la ventaja respecto al predecesor de que el ajuste va a ser mejor si las imágenes se encuentran alineadas; pero si por el contrario, las imágenes presentan un desalineamiento severo y no fueron registradas previamente, la estimación de la ganancia no va a efectuarse correctamente. Observe como la Ecuación 3.3 realiza la estimación teniendo en cuenta píxels en la misma posición de la imagen.

La Figura 3.1c muestra un ejemplo de aplicación de las técnicas de ajuste de ganancia sobre un par de imágenes. El tiempo de exposición y factor ISO para la imagen borrosa de la Figura 3.1a fue de 1/3 segundos y 200, respectivamente, mientras que la observación de corta exposición de la Figura 3.1b presenta un tiempo de exposición de 1/200 segundos e ISO 400. La apertura se mantuvo constante en ambas tomas con un valor de f/8. Usando estos parámetros en la Ec. 3.1 se obtuvo un valor $\lambda = 33,3$, cuya aplicación sobre la imagen de la Fig. 3.1b dió como resultado la imagen mostrada en la Fig. 3.1c. A la vista de los resultados, la imagen calibrada fotométricamente claramente sobrepasa el nivel de luminosidad de la imagen borrosa. La Figura 3.1d muestra el resultado de aplicar el parámetro de luminosidad obtenido mediante el uso de la Ecuación 3.3 que dió un valor $\lambda = 24,3$. Como se observa el parámetro se ajusta mejor que el método directo pero sigue sin ser perfecto por tratarse de una aproximación global. Nótese que en esta imagen la estimación es bastante buena ya que las imágenes de entrada no presentan una desalineación severa.

3.1.2. Ecuación de histograma mediante histograma de imagen borrosa

Este método es el que se ha utilizado para realizar el calibrado fotométrico de la imagen de baja luminosidad en todos los experimentos de la memoria. Para ello hacemos uso del histograma de la banda de luminancia (Y) de la imagen borrosa, que nos aporta valiosa información sobre la distribución que debería seguir una imagen con los niveles de luminosidad apropiados. El procedimiento de calibrado fotométrico se realiza mediante la rutina `histeq` que implementa MATLAB, esta rutina recibe como argumento el histograma de la imagen borrosa y devuelve la transformación T que minimiza la diferencia entre el histograma deseado, que pasamos como segundo argumento de la función `histeq`, y el histograma de la imagen de corta exposición, es decir,

$$T = \arg \min_T |c_1(T(\mathbf{y}_2)) - c_2(\mathbf{y}_2)|, \quad (3.4)$$

donde c_1 y c_2 representan los histogramas acumulados de la observación borrosa y oscura respectivamente. La imagen calibrada fotométricamente se obtiene de aplicar la transformación T sobre la imagen oscura inicial, es decir

$$\mathbf{y}_{2\text{calibrada}} = T(\mathbf{y}_2). \quad (3.5)$$

El resultado de aplicar la transformación calculada en las imágenes de entrada de las Figuras 3.2a y 3.2b se muestra en la Figura 3.2c. Los resultados en la Figura 3.2 exhiben un ajuste correcto; sin embargo el ruido aumenta proporcionalmente al incremento de luminosidad que debemos realizar sobre la imagen oscura. Un inconveniente añadido de la imagen de corta exposición, es que tiene una escasa gama de tonos de luminancia (véase el histograma de la Figura 3.2c) debida a la cuantificación. La transformación obtenida para el par de imágenes de la Figura 3.2, se representa en la gráfica de la Figura 3.3. A modo de ejemplo, esta transformación haría corresponder un valor de luminancia de entrada de 50 de la imagen de corta exposición, con el nivel de luminancia 200 de salida.

3.2. Calibración geométrica

En general es muy complicado captar dos imágenes perfectamente alineadas ya que incluso usando un trípode, podría introducirse un leve desplazamiento entre disparos consecutivos simplemente por el accionamiento del disparador. Es por ello que se hace necesario aplicar un método de registrado de imágenes antes de ejecutar el algoritmo de restauración deseado. El procedimiento de calibrado geométrico selecciona una de las imágenes como la imagen de referencia y calcula los parámetros que deberemos aplicar sobre la segunda imagen para hacerla coincidir con la de referencia. Como ya hemos expresado previamente, es recomendable realizar un ajuste de luminosidad previo al calibrado geométrico, ya que de no ser así, el método de registrado podría incurrir en errores en la estimación de los parámetros asociados ya que los métodos asumen una iluminación semejante entre las diferentes imágenes. En esta sección se explica el uso de dos métodos para corregir la desalineación entre las imágenes de manera automática. En el caso de registrado por emparejamiento de bloques (Sec. 3.2.1) estimamos un desplazamiento de las componentes horizontal y vertical; mientras que el método que combina Surf+RANSAC estima una matriz de homografía (Sec. 3.2.2). Cada método aporta una serie de ventajas e inconvenientes. Así, el método de emparejamiento de bloques es muy rápido y

puede aplicarse sin realizar obligatoriamente un calibrado fotométrico previo; sin embargo únicamente efectivo en escenas que presenten traslaciones en el plano de imagen. Por otro lado, el método basado en la matriz de homografía es más lento pero aplicable a un conjunto más amplio de transformaciones, incluidas rotaciones y cambios de escala además de traslación. La precisión del calibrado va a influir directamente en el proceso de restauración, como por ejemplo en la estimación de la función de emborronamiento. Una mala calibración geométrica puede ocasionar errores en el cálculo de la PSF.

3.2.1. Registrado por emparejamiento de bloques (*Block Matching*)

El objetivo de esta técnica es obtener el mejor desplazamiento horizontal y vertical, en términos de menor error, que aplicado a la imagen a alinear, la ajuste sobre la imagen de referencia. Para ello se usa el algoritmo de emparejamiento de bloques (*block matching*) [66] en el que se compara una zona de las imágenes aplicando todos los desplazamientos posibles dentro de una ventana de búsqueda y se devuelve el desplazamiento que incurra en un menor error. Este método es efectivo siempre que la desalineación entre las imágenes sea provocada por una traslación en el plano de imagen. Ya que suponemos que el desplazamiento es global para toda la imagen, se propone utilizar un único bloque que sea representativo de la imagen para estimar el desplazamiento. En concreto, se propone usar un bloque en el centro de la imagen y de tamaño mitad de la misma en horizontal y vertical.

La ventana de búsqueda debe ser suficientemente grande para captar el desplazamiento entre las imágenes. En la mayoría de los casos, escoger un desplazamiento máximo de un 10% del tamaño de la imagen suele proporcionar resultados satisfactorios aunque puede ajustarse dicho tamaño en función de las imágenes a tratar. Es decir, si la imagen tiene un tamaño 200x200, usando un desplazamiento del 10% de su tamaño, buscaremos en una ventana de $+ - 20$ píxeles alrededor del centro de la imagen.

Como medida del error entre los bloques de las imágenes \mathbf{y}_1 e \mathbf{y}_2 usaremos la suma de diferencias al cuadrado (*Sum of Squared Differences*, SSD) definida como

$$SSD(dx, dy) = \sum_{i,j \in \text{bloque}} (\mathbf{y}_1(i + dy, j + dx) - \mathbf{y}_2(i, j))^2$$

donde dx y dy son el desplazamiento a evaluar en la dirección x e y , respectivamente, y *bloque* representa el soporte del bloque de la imagen que se considera.

Por último, el algoritmo de búsqueda que se utiliza es la búsqueda exhaustiva (*Full Search*)

[66] con precisión de un píxel ya que nos asegura encontrar el mínimo error posible. Es conveniente notar que, aunque es importante realizar un emparejamiento de las dos imágenes, no es necesario un registrado con precisión subpíxel, y de hecho pequeños errores de desplazamiento no afectarán al método de restauración, ya que estos pequeños desplazamientos se compensan posteriormente de forma automática en la estimación de la función de emborronamiento.

El método propuesto se ha aplicado al par de imágenes de la Figura 3.2a y 3.2c. Se utilizó como referencia la imagen ruidosa después del calibrado fotométrico, y por tanto, los parámetros estimados se aplican sobre la imagen borrosa. Los resultados pueden observarse en la Figura 3.4. El desplazamiento calculado ha sido de -2 píxeles tanto en horizontal como en vertical. El resultado de aplicar este desplazamiento sobre la imagen borrosa sin calibrar mostrada en la Figura 3.4a se muestra en la Figura 3.4b. La Figura 3.4c muestra la imagen ruidosa, calibrada fotométricamente, para comparación.

3.2.2. Registrado mediante Homografía (Surf+RANSAC)

En esta sección abordamos el problema del calibrado geométrico con una técnica más robusta que considera otros tipos de transformaciones (rotación y escala) además de la traslación. Sin embargo este proceso es más lento y requiere de varias etapas con alta complejidad. En la primera de ellas se estiman una serie de puntos característicos entre ambas imágenes, sobre los que posteriormente se establecerá una correspondencia. Para ello utilizamos Surf [67], uno de los métodos más utilizados para la extracción de características. Hay que notar que algunos de los puntos característicos proporcionados por Surf en ambas imágenes no estarán en correspondencia, debido, por una parte, al ruido de la observación de corta exposición y por otra, al emborronamiento de la imagen correctamente expuesta. Por tanto, será necesario recurrir a un algoritmo robusto de estimación de la matriz de homografía que nos permita descartar aquellos píxeles que no estén en correspondencia entre ambas imágenes. Para seleccionar aquellos valores más representativos del modelo (que en nuestro caso va a ser una homografía o matriz de transformación) usaremos el método robusto RANSAC [68]. La matriz de homografía del modelo se va a calcular usando el algoritmo de transformación lineal directa (*Direct Linear Transformation*, DLT [61]). Una vez estimada la matriz de homografía, se aplica sobre la imagen que deseamos calibrar geométricamente para obtener el par de imágenes registrado.

Surf (*Speeded Up Robust Features*) es un detector de características locales robusto, propuesto por Herbert Bay et al. en 2006 [67]. En nuestro caso, lo usaremos para extraer puntos

de interés y emparejarlos. Surf es relativamente robusto a rotación, escala y ruido, por tanto suficientemente adecuado para nuestra tarea. Cada punto de interés detectado queda representado por un vector de características (descriptor), y la correspondencia entre los puntos de las imágenes se realiza comparando la distancia euclídea entre los vectores descriptores. Es importante hacer notar que para aplicar el método de calibrado geométrico que proponemos en esta sección, es fundamental aplicar el calibrado fotométrico previo, ya que Surf no es robusto frente a diferencias de iluminación. La justificación de usar Surf frente a otros métodos de extracción de características como SIFT[69, 70] se debe, fundamentalmente a dos factores, el primero de ellos es la velocidad, Surf es más rápido que SIFT, y el segundo factor se debe al reducido número de puntos en correspondencia que obtuvo SIFT al aplicarlo sobre el problema de calibrado en imágenes con distinta luminosidad.

El conjunto de parejas de puntos proporcionados por Surf sirven de entrada al algoritmo de estimación de la homografía. RANSAC («RANdom SAmple Consensus») es un método iterativo que permite estimar los parámetros de un modelo matemático, en nuestro caso la matriz de homografía, de forma robusta a partir de un conjunto de datos observados (puntos en correspondencia) que contienen *outliers* (valores atípicos o que no se ajustan al modelo). El algoritmo fue originalmente propuesto por Fischler y Bolles en 1981 [68]. Esta técnica la usaremos en conjunción con el algoritmo *Direct Linear Transformation*, (DLT) [61] que permite estimar los parámetros del modelo (homografía) que mejor se ajusta a los datos.

El algoritmo se compone de dos pasos clave en cada iteración: En el primero de ellos se escoge aleatoriamente un subconjunto mínimo de datos para construir un modelo de matriz de homografía usando el algoritmo DLT. El segundo paso se evalúa dicho modelo sobre el conjunto de datos global, conservando la mejor solución hasta el momento y utilizando el estancamiento como condición de parada, es decir que no se mejore esa solución en una serie de iteraciones. Partiendo del conjunto completo de muestras, es decir, de los puntos en correspondencia de un par de imágenes, se ejecuta el algoritmo al menos 1000 iteraciones, ajustando la homografía mediante todos los puntos que se ajusten al modelo (que se estiman en cada iteración) con un mínimo de 4 puntos. Esta restricción es debida a que el algoritmo DLT necesita un mínimo de 4 puntos para estimar la homografía. En nuestros ejemplos hemos fijado la proporción de puntos que se ajustan al modelo a 99%. El resultado de la aplicación del algoritmo es la matriz de homografía, el conjunto de muestras que se utilizó para estimarla (*consensus set*) y el error del modelo respecto a los datos.

En la Figura 3.5 se muestra un ejemplo de la aplicación de este algoritmo sobre el par de imágenes de la Figura 3.2. Nótese que es imprescindible que la imagen ruidosa haya sido calibrada fotométricamente con la borrosa por lo que la entrada al algoritmo serán las imágenes de las Figuras 3.2a y 3.2c. La Figura 3.5a muestra los puntos en correspondencia que obtuvo surf en ambas imágenes unidos por una línea. Como se puede observar hay pocas líneas atípicas que unen puntos que, claramente, no se corresponden. La imagen rectificadora de la Figura 3.5b se obtuvo después de aplicar la homografía sobre la imagen borrosa. Esta transformación claramente resuelve una leve rotación en sentido contrario a las agujas del reloj entre las imágenes. Después de la transformación, las dimensiones de la imagen rectificadora suelen diferir de la imagen de referencia, por ello recomendamos aplicar una reducción del tamaño de la imagen para eliminar el marco negro que se añade por el efecto de la rotación, al tamaño deseado desde el centro de las imágenes. La Figura 3.5c representa la imagen borrosa antes de realizar el registro, mientras que la Figura 3.5d muestra el resultado final y la Figura 3.5e muestra la imagen ruidosa. Estas tres imágenes han sido enmarcadas al mismo tamaño para facilitar la comparación visual y muestran el resultado del proceso de calibrado propuesto.

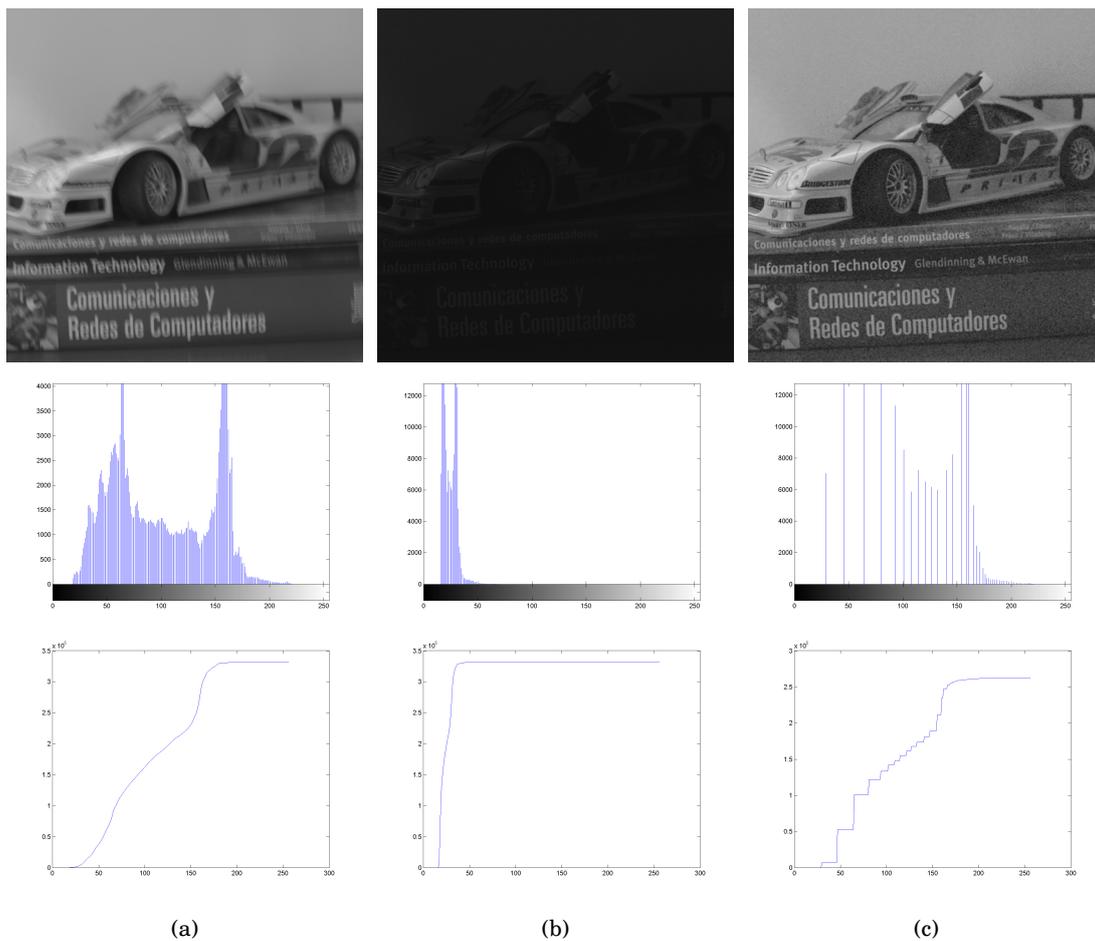


Figura 3.2: (a) Imagen borrosa, histograma e histograma acumulado. (b) Imagen de corta exposición, su histograma y el histograma acumulado. (c) Resultado de aplicar método de equalización de histograma con histograma de la imagen borrosa (a), sobre la imagen oscura (b), debajo se muestra su histograma y el histograma acumulado.

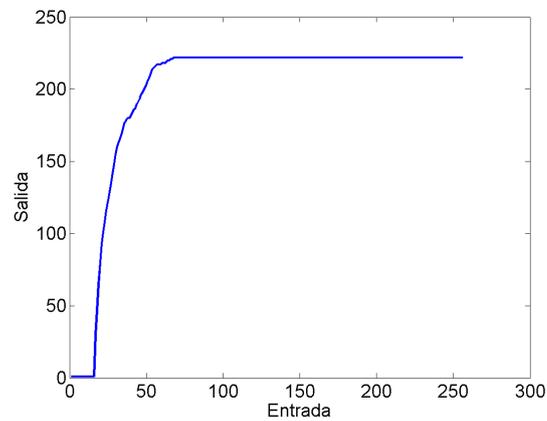


Figura 3.3: Transformación obtenida para cada uno de los «bins», en el eje de abscisas, con su correspondiente valor de gris de salida en ordenadas.

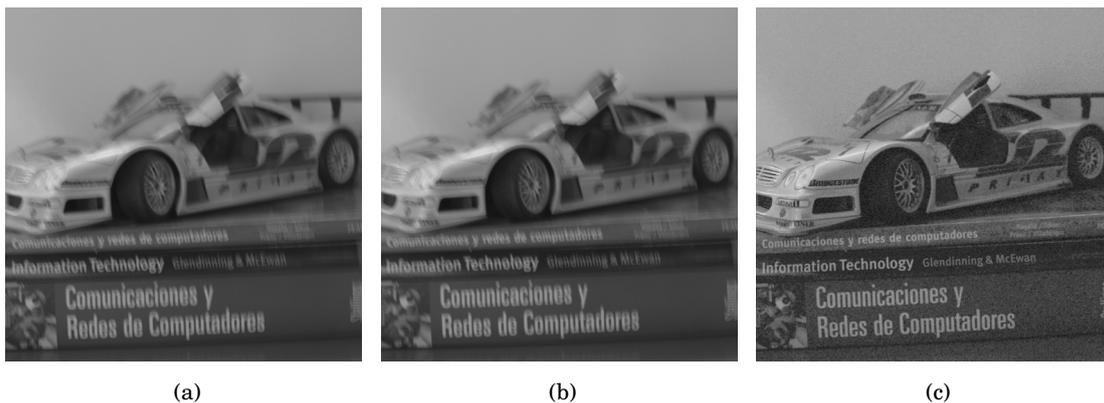


Figura 3.4: (a) imagen borrosa sin calibrar, (b) Imagen borrosa calibrada mediante emparejamiento de bloques, con un desplazamiento calculado de $(-2, -2)$ en horizontal y vertical, respectivamente. (c) imagen ruidosa calibrada fotométricamente.

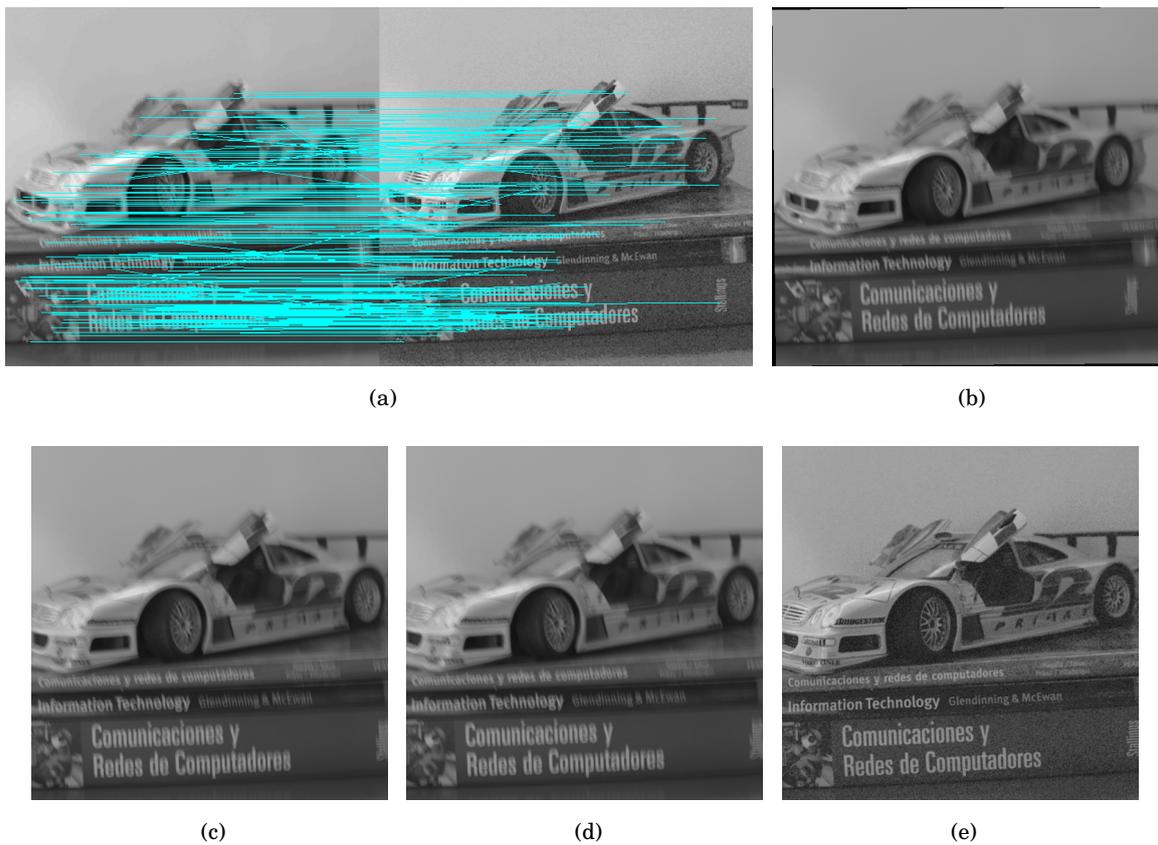


Figura 3.5: (a) imágenes de entrada con descriptores Surf. Cada línea une puntos en correspondencias. (b) Imagen borrosa rectificada (con la homografía aplicada). (c) Imagen borrosa (enmarcada) sin calibrado geométrico. (d) Imagen borrosa calibrada y enmarcada. (e) Imagen ruidosa enmarcada y calibrada fotométricamente para comparación visual del registrado.

Capítulo 4

Deconvolución usando un núcleo espacialmente invariante

*Se quiere más aquello que se ha conseguido con
muchas fatigas*

Aristóteles (384 a.C.-322 a.C.).
Filósofo, lógico y científico griego.

4.1. Introducción

En este capítulo proponemos un método para restauración de un par de imágenes asumiendo un modelo de degradación espacialmente invariante. Para ello proponemos modelizar la degradación sufrida en cada una de las imágenes del par por separado para después realizar, en la inferencia bayesiana variacional, una combinación de los modelos de observación.

La combinación de los modelos de observación, al igual que la combinación de los modelos a priori de imagen, conlleva un gran desafío y representa un problema muy interesante. Sin embargo, mientras que la combinación de conocimiento a priori sobre la imagen ha recibido algún interés en la literatura reciente, véase, por ejemplo [71] y las referencias incluidas en éste, existen pocos trabajos en relación a la combinación de los modelos de observación, aunque de notable importancia el trabajo en [2] que propone un único modelo de observación pero que añade un término combinación de las dos observaciones.

En este capítulo se ha desarrollado un método de restauración a partir de pares de imáge-

nes, asumiendo un emborronamiento espacialmente invariante, que permite combinar modelos de observación. En concreto, definimos dos modelos de observación combinación de tres observaciones: la observación borrosa, la ruidosa (ambas previamente calibradas), y una tercera observación obtenida de la diferencia entre la imagen borrosa y una versión borrosa de la imagen ruidosa obtenida por convolución con el emborronamiento estimado. Cada uno de estos dos modelos de observación se combinará con la información a priori y, usando el paradigma Bayesiano, podríamos obtener dos distribuciones a posteriori diferentes, una por cada combinación de modelo de observación y modelos a priori. Sin embargo, es deseable obtener una única distribución a posteriori con la combinación de toda la información disponible. El método propuesto se fundamenta en el uso de la divergencia de Kullback-Leibler y la teoría variacional para la aproximación de la distribución a posteriori resultado de minimizar una combinación lineal convexa de la divergencia de Kullback-Leibler asociada a cada distribución a posteriori. De esta forma, somos capaces de encontrar la distribución de forma explícita.

El resto del capítulo se organiza de la siguiente forma: En la Sección 4.2 formulamos el problema de combinación de pares de imágenes borrosa/ruidosa e introducimos matemáticamente el tercer modelo de observación. En la Sección 4.3 se modelan las variables desconocidas mediante un marco Bayesiano jerárquico y presentamos la distribución a posteriori asociada a cada combinación de los modelos de observación. Aplicando inferencia variacional, en la Sec. 4.4 encontramos una única aproximación a la distribución a posteriori que incorpora la información de los tres modelos de observación. Para concluir el presente capítulo, en la Sec. 4.6 analizamos el funcionamiento del método propuesto mediante una serie de experimentos.

4.2. Formulación del problema

Asumiendo que el emborronamiento se debe principalmente al movimiento de la cámara durante el tiempo de exposición normal, que modelamos como un operador lineal espacialmente invariante, y que las dos imágenes están calibradas fotométricamente y geométricamente con antelación, el proceso de formación de las dos imágenes que forman el par de imágenes puede ser escrito como

$$\mathbf{y}_1 = \mathbf{H}\mathbf{x} + \mathbf{n}_1, \quad (4.1)$$

$$\mathbf{y}_2 = \mathbf{x} + \mathbf{n}_2, \quad (4.2)$$

donde \mathbf{y}_1 e \mathbf{y}_2 representan, respectivamente, la observación borrosa y ruidosa, \mathbf{x} la imagen original desconocida, y \mathbf{n}_1 y \mathbf{n}_2 representan el ruido que asumimos que es ruido blanco Gaussiano de media cero y varianzas β_1^{-1} y β_2^{-1} , respectivamente. Es decir,

$$\mathbf{n}_1 \sim \mathcal{N}(\mathbf{H}\mathbf{x}, \beta_1^{-1}\mathbf{I}), \quad (4.3)$$

$$\mathbf{n}_2 \sim \mathcal{N}(\mathbf{x}, \beta_2^{-1}\mathbf{I}). \quad (4.4)$$

Nótese que $\beta_1 \gg \beta_2$, ya que la imagen \mathbf{y}_1 está correctamente expuesta, mientras \mathbf{y}_2 posee un alto nivel de ruido. Usando la notación vectorial, \mathbf{y}_1 , \mathbf{y}_2 , \mathbf{x} , \mathbf{n}_1 y \mathbf{n}_2 denotan vectores de dimensión $N \times 1$, donde N indica el número de píxeles de las imágenes. La matriz de convolución \mathbf{H} de dimensiones $N \times N$ modela la PSF, \mathbf{h} , de tamaño M , con $M \leq N$.

Debido a que \mathbf{y}_1 e \mathbf{y}_2 provienen de la misma escena, se encuentran altamente correladas. Este hecho permite definir un modelo de degradación adicional como en [2], el cual aprovecha la dependencia entre las observaciones \mathbf{y}_1 e \mathbf{y}_2 , con una única incógnita \mathbf{H} . Este tercer modelo además de relacionar las observaciones a partir del emborronamiento, cumple la condición de coprimalidad por el hecho que el único factor común es una constante. El uso de esta condición de coprimalidad se ha utilizado en otros métodos tales como en [72]. La combinación de los modelos de observación en las Ecs. (4.1) y (4.2) dado \mathbf{H} que se propone es

$$\mathbf{y}_1 - \mathbf{H}\mathbf{y}_2 = \mathbf{n}_{12}. \quad (4.5)$$

Teniendo en cuenta las Ecuaciones 4.3 y 4.4, obtenemos que

$$(\mathbf{y}_1 - \mathbf{H}\mathbf{y}_2 | \mathbf{h}, \mathbf{y}_1, \mathbf{y}_2) \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{n}_{12}}, \boldsymbol{\Sigma}_{\mathbf{n}_{12}}), \quad (4.6)$$

donde

$$\boldsymbol{\mu}_{\mathbf{n}_{12}} = \langle \mathbf{n}_1 - \mathbf{H}\mathbf{n}_2 \rangle = 0, \quad (4.7)$$

$$\boldsymbol{\Sigma}_{\mathbf{n}_{12}} = \langle (\mathbf{n}_1 - \mathbf{H}\mathbf{n}_2)(\mathbf{n}_1 - \mathbf{H}\mathbf{n}_2)^t \rangle = \beta_1\mathbf{I} + \beta_2\mathbf{H}\mathbf{H}^t. \quad (4.8)$$

Considerando ruido incorrelado, la matriz de covarianzas se puede aproximar por una matriz diagonal con los elementos

$$\beta_{12}^{-1} = \beta_1^{-1} + \beta_2^{-1} \sum_i h_i^2,$$

con $i = 1, \dots, N$ en la diagonal. Usando esta aproximación,

$$|\beta_1\mathbf{I} + \beta_2\mathbf{H}\mathbf{H}^t|^{-1/2} \simeq \beta_{12}^{-N/2},$$

con $\beta_{12} > 0$ y, por tanto, en el resto del capítulo consideraremos que

$$(\mathbf{y}_1 - \mathbf{H}\mathbf{y}_2 | \mathbf{h}, \mathbf{y}_1, \mathbf{y}_2) \sim \mathcal{N}(0, \beta_{12}^{-1} \mathbf{I}).$$

El objetivo del problema de restauración del par borrosa/ruidosa consiste en la estimación de la imagen nítida \mathbf{x} y emborronamiento \mathbf{h} utilizando \mathbf{y}_1 e \mathbf{y}_2 , junto al conocimiento a priori sobre \mathbf{x} , \mathbf{h} , \mathbf{n}_1 y \mathbf{n}_2 .

4.3. Modelo Bayesiano jerárquico

Como se describió en el Capítulo 2 en esta memoria, adoptamos el marco Bayesiano jerárquico que consta de dos etapas. En la primera etapa, se definen las distribuciones a priori de la imagen desconocida, \mathbf{x} , y la PSF, \mathbf{h} , y se proponen distribuciones de probabilidad para los modelos de degradación descritos en la sección anterior. Las distribuciones definidas en la primera etapa dependerán de ciertos parámetros, llamados «hiperparámetros», que están modelados por distribuciones a priori en la segunda etapa. En la Figura 4.1 se muestra el diagrama que relaciona las variables conocidas y desconocidas del modelo jerárquico. A continuación procedemos a describir las distribuciones de probabilidad que usaremos.

4.3.1. Primera etapa: Modelos a priori

Dado que el emborronamiento es causado principalmente por el movimiento de la cámara durante el tiempo de exposición, éste presenta las características típicas del desenfoque por movimiento uniforme. Por lo tanto, se espera que tenga una estructura rala, es decir, la mayoría de los coeficientes de la PSF van a ser cero o muy pequeños. Con el fin de aprovechar esta información, utilizamos una mezcla de D distribuciones exponenciales como modelo a priori para cada coeficiente de la PSF, es decir,

$$p(\mathbf{h} | \{\tau_d\}, \{\sigma_d\}) = \prod_{j=1}^M \left(\sum_{d=1}^D \tau_d \text{Expon}(h_j | \sigma_d) \right) \quad (4.9)$$

con τ_d el coeficiente de la mezcla para la clase d y σ_d el parámetro de cada distribución exponencial, y Expon corresponde a la función exponencial definida como

$$\text{Expon}(h_j | \sigma_d) = \begin{cases} \sigma_d \exp(-\sigma_d h_j) & \text{si } h_j \geq 0, \\ 0 & \text{si } h_j < 0. \end{cases}$$

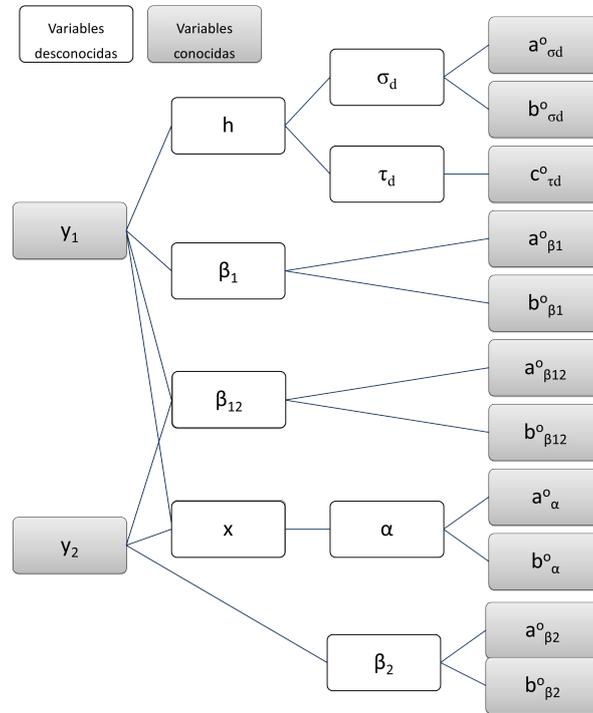


Figura 4.1: Modelo jerárquico propuesto.

Además de imponer raleza, el conocimiento a priori también implica positividad en los coeficientes de la PSF h_j , este hecho es especialmente útil ya que no imponemos esa restricción de forma artificial; como si ocurre, por ejemplo, en [65, 73]. Nótese que esta mezcla de exponenciales como a priori también se ha utilizado antes para modelar PSFs típicas del movimiento de la cámara [26] así como en el análisis de componentes independientes [51].

Como modelo a priori de imagen se utiliza la función de variación total (TV) ya que conserva las fronteras de los objetos en la imagen, sin sobre-ajustar, imponiendo suavidad al mismo tiempo [74]. Así que, podemos escribir la distribución a priori de la imagen como

$$p(\mathbf{x}|\alpha) \propto \frac{1}{Z_{\text{TV}(\alpha)}} \exp[-\alpha \text{TV}(\mathbf{x})], \quad (4.10)$$

donde Z_{TV} es la función de partición y $\text{TV}(\mathbf{x})$ se define como

$$\text{TV}(\mathbf{x}) = \sum_{i=1}^N \sqrt{(\Delta_i^h(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2}, \quad (4.11)$$

con los operadores $\Delta_i^h(\mathbf{x})$ y $\Delta_i^v(\mathbf{x})$ correspondientes a las derivadas horizontales y verticales de primer orden, en cada píxel i , respectivamente, es decir, $\Delta_i^h(\mathbf{x}) = x_i - x_{1e(i)}$ y $\Delta_i^v(\mathbf{x}) = x_i -$

$x_{\text{up}(i)}$, donde $\text{le}(i)$ y $\text{up}(i)$ representan los vecinos más cercanos a i , a la izquierda y arriba, respectivamente. La función de partición Z_{TV} podemos aproximarla teniendo en cuenta que [75]

$$\int_r \int_s \exp \left[-\alpha \sqrt{r^2 + s^2} \right] dr ds = 2\pi/\alpha^2, \quad (4.12)$$

quedando finalmente el modelo a priori sobre la imagen como

$$p(\mathbf{x}|\alpha) \propto c \alpha^{N/2} \exp \left[-\alpha \sum_{i=1}^N \sqrt{(\Delta_i^h(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2} \right], \quad (4.13)$$

donde c es una constante y N corresponde al tamaño de la imagen.

4.3.2. Primera etapa: Modelo de observación

A partir de los modelos de degradación en las Ecuaciones (4.1), (4.2) y (4.5) definimos dos modelos de observación. El primero de ellos combinación de las Ecuaciones (4.1) y (4.2)

$$p_1(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \beta_1, \beta_2) \propto \beta_1^{N/2} \beta_2^{N/2} \exp \left[-\frac{\beta_1}{2} \|\mathbf{y}_1 - \mathbf{H}\mathbf{x}\|^2 - \frac{\beta_2}{2} \|\mathbf{y}_2 - \mathbf{x}\|^2 \right], \quad (4.14)$$

y el segundo modelo de observación, combinación de las Ecs. (4.2) y (4.5) expresado como

$$p_2(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \beta_2, \beta_{12}) \propto \beta_2^{N/2} \beta_{12}^{N/2} \exp \left[-\frac{\beta_{12}}{2} \|\mathbf{y}_1 - \mathbf{H}\mathbf{y}_2\|^2 - \frac{\beta_2}{2} \|\mathbf{y}_2 - \mathbf{x}\|^2 \right]. \quad (4.15)$$

En esta memoria nos centraremos en los dos modelos anteriores, si bien la extensión de esta teoría a otros modelos de observación alternativos es directa.

Nótese que en un principio podríamos haber considerado un único modelo de observación, como en [2] combinando las tres formas cuadráticas,

$$p(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \beta_1, \beta_2, \beta_{12}) \propto \frac{1}{Z(\beta_1, \beta_2, \beta_{12})} \exp \left[-\frac{\beta_1}{2} \|\mathbf{y}_1 - \mathbf{H}\mathbf{x}\|^2 - \frac{\beta_2}{2} \|\mathbf{y}_2 - \mathbf{x}\|^2 - \frac{\beta_{12}}{2} \|\mathbf{y}_1 - \mathbf{H}\mathbf{y}_2\|^2 \right], \quad (4.16)$$

sin embargo la estimación de los parámetros es complicada para esta combinación y requiere calcular la función de partición Z . En contraposición, los modelos propuestos nos permitirán estudiar teóricamente la forma de realizar la combinación y, además, determinar la mejor combinación de ambos modelos de observación, como evidenciamos en las siguientes secciones.

4.3.3. Segunda etapa: Distribuciones sobre los hiperparámetros

Los modelos de observación y a priori dependen de un conjunto de parámetros cuyo valor es crucial para determinar el rendimiento del algoritmo. Para modelizar la incertidumbre en los parámetros vamos a emplear distribuciones Gamma en α , β_1 , β_2 , β_{12} y σ_d , con $d = 1, \dots, D$, es decir,

$$p(\xi) = \text{Gamma}(\xi | a_\xi^o, b_\xi^o), \quad (4.17)$$

donde ξ denota un hiperparámetro con $a_\xi^o > 0$ y $b_\xi^o > 0$ los parámetros de forma e inversa de la escala de la distribución Gamma, con la distribución $\text{Gamma}(\xi | a_\xi^o, b_\xi^o)$ definida como

$$\Gamma(\xi | a_\xi^o, b_\xi^o) = \frac{(b_\xi^o)^{a_\xi^o}}{\Gamma(a_\xi^o)} \xi^{a_\xi^o - 1} \exp[-\xi b_\xi^o]. \quad (4.18)$$

La distribución Gamma tiene las siguientes propiedades

$$\mathbf{E}[\xi] = a_\xi^o / b_\xi^o, \quad \mathbf{Var}[\xi] = a_\xi^o / b_\xi^o, \quad \mathbf{Mode}[\xi] = (a_\xi^o - 1) / b_\xi^o. \quad (4.19)$$

La razón principal por la que elegimos distribuciones Gamma para modelar los hiperparámetros, es que son conjugadas con la distribución Gaussiana y, por tanto, las distribuciones a posteriori seguirán siendo distribuciones Gamma, facilitándonos la derivación de la aproximación de la distribución a posteriori considerablemente.

Para modelar la mezcla τ_d hacemos uso de la distribución Dirichlet con parámetros $c_{\tau_d}^o$

$$p(\{\tau_d\}_{d=1}^D) = \text{Dirichlet}(\{\tau_d\}_{d=1}^D | \{c_{\tau_d}^o\}_{d=1}^D), \quad (4.20)$$

con la distribución de Dirichlet definida por

$$\text{Dirichlet}(\{\tau_d\}_{d=1}^D | \{c_{\tau_d}^o\}_{d=1}^D) = \frac{\Gamma(c_{\tau_0}^o)}{\Gamma(c_{\tau_1}^o) \dots \Gamma(c_{\tau_D}^o)} \prod_D \tau_d^{(c_{\tau_d}^o - 1)}. \quad (4.21)$$

Las propiedades de esta distribución nos van a ayudar en el proceso de estimación ya que, por definición, $\sum_d^D \tau_d = 1$ y $0 \leq \tau_d \leq 1$, es decir, va a asegurar que los valores del parámetro de mezcla de cada clase d tome valores positivos y la suma de los parámetros de mezcla para todas las clases es 1, lo que a su vez asegura que la mezcla de varias distribuciones de probabilidad sigue siendo una distribución de probabilidad.

Finalmente, combinando (4.13), (4.17) y (4.20) con los dos modelos de observación (4.14) y

(4.15) obtenemos dos distribuciones conjuntas, $p_1(\cdot)$ y $p_2(\cdot)$, que vienen dadas por

$$p_1(\mathbf{y}_1, \mathbf{y}_2, \Omega, \boldsymbol{\beta}_1) = p(\mathbf{x}|\alpha)p(\alpha) \times p(\mathbf{h}|\{\tau_d\}, \{\sigma_d\}) \prod_{d=1}^D p(\tau_d)p(\sigma_d) \times p_1(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}, \mathbf{h}, \boldsymbol{\beta}_1)p(\boldsymbol{\beta}_1), \quad (4.22)$$

$$p_2(\mathbf{y}_1, \mathbf{y}_2, \Omega, \boldsymbol{\beta}_2) = p(\mathbf{x}|\alpha)p(\alpha) \times p(\mathbf{h}|\{\tau_d\}, \{\sigma_d\}) \prod_{d=1}^D p(\tau_d)p(\sigma_d) \times p_2(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}, \mathbf{h}, \boldsymbol{\beta}_2)p(\boldsymbol{\beta}_2), \quad (4.23)$$

donde $\Omega = \{\mathbf{x}, \mathbf{h}, \alpha, \{\tau_d\}, \{\sigma_d\}\}$, $\boldsymbol{\beta}_1 = \{\beta_1, \beta_2\}$ y $\boldsymbol{\beta}_2 = \{\beta_2, \beta_{12}\}$.

4.4. Inferencia Bayesiana Variacional

En la formulación Bayesiana la inferencia se basa en la distribución a posteriori. Por desgracia en nuestro caso ésta no puede calcularse analíticamente por varias razones, en primer lugar la dimensión del espacio de búsqueda de las posibles soluciones, en segundo lugar debido a la expresión compleja de la distribución a posteriori cuyas esperanzas son intratables. Por tanto, en esta memoria haremos uso de la teoría variacional para aproximarla. Llamaremos Θ al conjunto de incógnitas, es decir, $\Theta = \{\Omega, \boldsymbol{\beta}\}$ con $\boldsymbol{\beta} = \{\beta_1, \beta_2, \beta_{12}\}$. El objetivo consiste en aproximar la distribución a posteriori $p(\Theta|\mathbf{y}_1, \mathbf{y}_2)$ por otra distribución $q(\Theta)$ que nos permita realizar el análisis. En esta memoria proponemos aproximar la distribución a posteriori $p(\Theta|\mathbf{y}_1, \mathbf{y}_2)$ por la distribución que minimiza la siguiente combinación lineal convexa de medidas de divergencia de Kullback-Leibler (KL)

$$\hat{q}(\Theta) = \operatorname{argmin}_{q(\Theta)} \sum_{l=1}^2 \lambda_l C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| p_l(\Omega, \boldsymbol{\beta}_l|\mathbf{y}_1, \mathbf{y}_2)), \quad (4.24)$$

con $l = 1, 2$, $\lambda_l \geq 0$, $\lambda_1 + \lambda_2 = 1$, cuya divergencia de KL viene dada por

$$\begin{aligned} C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| p_l(\Omega, \boldsymbol{\beta}_l|\mathbf{y}_1, \mathbf{y}_2)) &= \int q(\Omega)q(\boldsymbol{\beta}_l) \log \left(\frac{q(\Omega)q(\boldsymbol{\beta}_l)}{p_l(\Omega, \boldsymbol{\beta}_l|\mathbf{y}_1, \mathbf{y}_2)} \right) d\Omega d\boldsymbol{\beta}_l. \\ &= \int q(\Omega)q(\boldsymbol{\beta}_l) \log \left(\frac{q(\Omega)q(\boldsymbol{\beta}_l)}{p_l(\mathbf{y}_1, \mathbf{y}_2, \Omega, \boldsymbol{\beta}_l)} \right) d\Omega d\boldsymbol{\beta}_l + \text{const} \end{aligned} \quad (4.25)$$

que siempre será positiva y será cero únicamente cuando $q(\Omega)q(\boldsymbol{\beta}_l) = p_l(\Omega, \boldsymbol{\beta}_l|\mathbf{y}_1, \mathbf{y}_2)$. Generalmente se asume [54] que la aproximación a la distribución se puede factorizar, utilizando la

factorización siguiente

$$q(\Omega) = q(\mathbf{x})q(\mathbf{h})q(\alpha)q(\{\tau_d\}_{d=1}^D) \prod_{d=1}^D q(\sigma_d),$$

$$q(\boldsymbol{\beta}_1) = q(\beta_1)q(\beta_2),$$

$$q(\boldsymbol{\beta}_2) = q(\beta_2)q(\beta_{12}),$$

$$q(\mathbf{h}) = \prod_{j=1}^M q(h_j),$$

$$q(\Theta) = q(\Omega)q(\beta_1)q(\beta_2)q(\beta_{12}).$$

Esta forma de factorización de la inferencia variacional coincide con otro marco desarrollado en física conocido como *mean field theory* (teoría del campo medio) [76]. La estimación de λ_1 y λ_2 no se tratará en esta memoria pero demostraremos empíricamente que una combinación no degenerada de divergencias, es decir, $\lambda_1, \lambda_2 > 0$, proporciona mejores resultados que una degenerada. Nótese que

$$\int q(\Omega)q(\boldsymbol{\beta}_l) \log \left(\frac{q(\Omega)q(\boldsymbol{\beta}_l)}{p_l(\mathbf{y}_1, \mathbf{y}_2, \Omega, \boldsymbol{\beta}_l)} \right) d\Omega d\boldsymbol{\beta}_l = \int q(\Theta) \log \left(\frac{q(\Omega)q(\boldsymbol{\beta}_l)}{p_l(\mathbf{y}_1, \mathbf{y}_2, \Omega, \boldsymbol{\beta}_l)} \right) d\Theta, \quad (4.26)$$

de esta manera la expresión (4.24) puede reescribirse en forma más compacta como

$$\hat{q}(\Theta) = \operatorname{argmin}_{q(\Theta)} \int q(\Theta) \log \left(\frac{q(\Omega)}{p(\Omega)} \prod_{l=1}^2 \left[\frac{q(\boldsymbol{\beta}_l)}{p_l(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \boldsymbol{\beta}_l) p(\boldsymbol{\beta}_l)} \right]^{\lambda_l} \right) d\Theta, \quad (4.27)$$

con $p(\Omega) = p(\mathbf{x}|\alpha)p(\alpha) \times p(\mathbf{h}|\{\tau_d\}, \{\sigma_d\}) \prod_{d=1}^D p(\tau_d)p(\sigma_d)$. Queremos hacer notar que el modelo en [2] corresponde a escoger $\lambda_1 = \lambda_2 = 1$.

Lamentablemente, los resultados generales del análisis Bayesiano Variacional no pueden ser directamente utilizados en este caso, ya que el modelo a priori TV y el de mezcla hacen que sea inviable el cálculo de la divergencia de KL en (4.25). Los problemas causados por el modelo TV pueden evitarse mediante la utilización de un enfoque de minimización/mayorización, cuyos detalles se pueden consultar en [74], el cual permite acotar la distribución en (4.13), haciendo posible la derivación analítica de la inferencia bayesiana. A continuación exponemos un resumen de estos resultados.

A partir de la desigualdad de las medias aritmética/geométrica, que establece que la media aritmética de un conjunto de números reales positivos es mayor o igual que la media geométrica del mismo conjunto, cumpliéndose la igualdad cuando todos los elementos del conjunto son iguales, obtenemos que para cualquier par de números reales $u \geq 0$ y $v \geq 0$

$$\sqrt{uv} \leq \frac{u+v}{2} \Rightarrow \sqrt{u} \leq \frac{u+v}{2\sqrt{v}}, \quad (4.28)$$

Usando esta desigualdad en la Ecuación (4.13), definimos el funcional

$$\mathbf{M}(\alpha, \mathbf{x}, \mathbf{w}) = c \alpha^{N/2} \exp \left[-\frac{\alpha}{2} \sum_{i=1}^N \frac{(\Delta_i^h(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2 + w_i}{\sqrt{w_i}} \right], \quad (4.29)$$

donde hemos usado que $u = (\Delta_i^h(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2$ y $v = w_i$ con $\mathbf{w} \in (\mathfrak{R}^+)^N$ el vector N -dimensional de componentes $w_i, i = 1, \dots, N$ que necesitará calcularse y posee, como veremos más adelante, una interpretación intuitiva relacionada con la imagen desconocida \mathbf{x} . Comparando (4.29) con (4.13) puede verse que el funcional $\mathbf{M}(\alpha, \mathbf{x}, \mathbf{w})$ es una cota inferior de la a priori de la imagen $p(\mathbf{x}|\alpha)$, es decir,

$$p(\mathbf{x}|\alpha) \geq \mathbf{M}(\alpha, \mathbf{x}, \mathbf{w}). \quad (4.30)$$

Usando este resultado, podemos encontrar una cota inferior de la distribución conjunta en (4.22) y (4.23) como,

$$\begin{aligned} p_l(\mathbf{y}_1, \mathbf{y}_2, \Omega, \boldsymbol{\beta}_l) &\geq \mathbf{M}(\alpha, \mathbf{x}, \mathbf{w}) p(\alpha) p(\mathbf{h}|\{\tau_d\}, \{\sigma_d\}) \prod_{d=1}^D p(\tau_d) p(\sigma_d) p_l(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \boldsymbol{\beta}_l) p(\boldsymbol{\beta}_l) \\ &= \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2), \end{aligned} \quad (4.31)$$

para $l = 1, 2$, la cual lleva a la siguiente cota superior para la divergencia de KL en (4.25)

$$C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| p_l(\Omega, \boldsymbol{\beta}_l | \mathbf{y}_1, \mathbf{y}_2)) \leq C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2)) + \text{const.} \quad (4.32)$$

Es posible ajustar la cota superior $C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2))$ minimizándola respecto a \mathbf{w} , ya que

$$C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| p_l(\Omega, \boldsymbol{\beta}_l | \mathbf{y}_1, \mathbf{y}_2)) \leq \min_{\mathbf{w}} C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2)) + \text{const.} \quad (4.33)$$

Por lo tanto, al minimizar la cota superior $C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2))$ respecto a $q(\Omega), q(\boldsymbol{\beta}_l)$ y \mathbf{w} , reducimos iterativamente la distancia entre la cota superior y la divergencia $C_{KL}(q(\Omega)q(\boldsymbol{\beta}_l) \| p_l(\Omega, \boldsymbol{\beta}_l | \mathbf{y}_1, \mathbf{y}_2))$. Este hecho hace que podamos utilizar la cota superior como aproximación a la distancia KL original, y el análisis Bayesiano Variacional se puede efectuar usando esta cota en su lugar.

Para cada incógnita $\theta \in \Theta$, la aproximación a la distribución $q(\theta)$ se lleva a cabo alternativamente mediante la minimización de la divergencia respecto a cada $q(\theta)$ manteniendo $q(\Theta_\theta)$ constante, donde Θ_θ denota el conjunto de incógnitas Θ excepto θ . Este enfoque resulta en la siguiente solución general [54]

$$q(\theta) = \text{const} \times \exp(\mathbf{E}_{q(\Theta_\theta)} [\log \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2)]) \quad (4.34)$$

donde $\mathbf{E}_{q(\Theta_\theta)}$ expresa la esperanza respecto a la distribución $q(\Theta_\theta)$. Con el fin de resolver (4.34), es necesaria una aproximación adicional a la Ecuación (4.9) por utilizar una mezcla de distribuciones a priori. Concretamente utilizamos la inecuación de Jensen de forma similar a [51]. La inecuación de Jensen dice que dada una función cóncava φ , una serie de números c_1, \dots, c_n en su dominio y de pesos d_i se cumple que

$$\varphi\left(\frac{\sum_i d_i c_i}{\sum_i d_i}\right) \geq \frac{\sum_i d_i \varphi(c_i)}{\sum_i d_i}. \quad (4.35)$$

La aplicación directa de este resultado nos permite acotar el logaritmo de una sumatoria por la sumatoria de logaritmos de la siguiente forma

$$\log(p(\mathbf{h}|\{\tau_d\}, \{\sigma_d\})) = \log\left[\prod_{j=1}^M \left(\sum_{d=1}^D \tau_d \text{Expon}(h_j | \sigma_d)\right)\right] \geq \sum_{j=1}^M \sum_{d=1}^D \mu_{jd} \log\left(\frac{\tau_d}{\mu_{jd}} \text{Expon}(h_j | \sigma_d)\right), \quad (4.36)$$

con $\mu_{jd} \geq 0$, $\sum_{d=1}^D \mu_{jd} = 1$, $j = 1, \dots, M$. Podemos encontrar un análisis sobre el ajuste de esta cota en [51]. Las variables auxiliares μ_{jd} deben ser calculadas junto al resto de incógnitas Θ , como mostraremos ulteriormente.

Haciendo uso de la Ecuación (4.36) obtenemos una cota inferior para $\log \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2)$ de la siguiente forma:

$$\begin{aligned} \log \mathbf{F}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \mathbf{y}_1, \mathbf{y}_2) &= \log \mathbf{M}(\alpha, \mathbf{x}, \mathbf{w}) + \log p(\alpha) + \sum_{d=1}^D \log p(\tau_d) p(\sigma_d) + \log p(\mathbf{h}|\{\tau_d\}, \{\sigma_d\}) \\ &\quad + \log p_l(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \boldsymbol{\beta}_l) + \log p(\boldsymbol{\beta}_l) \\ &\geq \log \mathbf{M}(\alpha, \mathbf{x}, \mathbf{w}) + \log p(\alpha) + \sum_{j=1}^M \sum_{d=1}^D \mu_{jd} \log\left(\frac{\tau_d}{\mu_{jd}} \text{Expon}(h_j | \sigma_d)\right) \\ &\quad + \sum_{d=1}^D \log[p(\tau_d) p(\sigma_d)] + \log p_l(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \boldsymbol{\beta}_l) + \log p(\boldsymbol{\beta}_l) \\ &= \mathbf{B}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \boldsymbol{\mu}, \mathbf{y}_1, \mathbf{y}_2), \end{aligned} \quad (4.37)$$

con $\boldsymbol{\mu} = \{\mu_{jd} | j = 1, \dots, M, d = 1, \dots, D\}$.

Utilizando esta cota inferior, obtenemos las soluciones

$$q(\beta_1) = \text{const} \times \exp(\langle \mathbf{B}_1(\Omega, \boldsymbol{\beta}_1, \mathbf{w}, \boldsymbol{\mu}, \mathbf{y}_1, \mathbf{y}_2) \rangle_\Omega), \quad (4.38)$$

$$q(\beta_{12}) = \text{const} \times \exp(\langle \mathbf{B}_2(\Omega, \boldsymbol{\beta}_2, \mathbf{w}, \boldsymbol{\mu}, \mathbf{y}_1, \mathbf{y}_2) \rangle_\Omega), \quad (4.39)$$

donde $\langle \cdot \rangle_\Omega = \mathbf{E}_{q(\Omega)}[\cdot]$, y $\mathbf{E}_{q(\Omega)}$ denota la esperanza respecto la distribución $q(\Omega)$. Además, para calcular el resto de distribuciones, $q(\gamma)$, $\gamma \in \{\Omega, \beta_2\}$, debemos tener en cuenta ambas divergen-

cias, obteniendo

$$q(\gamma) = \text{const} \times \exp \left(\left\langle \sum_{l=1}^2 \lambda_l \mathbf{B}_l(\Omega, \boldsymbol{\beta}_l, \mathbf{w}, \boldsymbol{\mu}, \mathbf{y}_1, \mathbf{y}_2) \right\rangle_{\Theta_\gamma} \right), \quad (4.40)$$

donde Θ_γ denota el conjunto de incógnitas sin γ .

El cálculo de cada una de las distribuciones para cada incógnita, determina un procedimiento iterativo, que converge a la mejor aproximación de la distribución a posteriori verdadera $p(\Theta|\mathbf{y}_1, \mathbf{y}_2)$ por distribuciones en la forma de (4.24). En la presente memoria, utilizamos la media de estas distribuciones como estimadores de las incógnitas. A continuación procedemos a hacer explícita la forma de cada una de estas distribuciones.

4.5. Cálculo de las aproximaciones de la distribución a posteriori

En esta sección, mostramos las formas explícitas de cada una de las distribuciones $q(\Theta)$. La distribución $q(\mathbf{x})$ se calcula a partir de (4.40) como una distribución Gaussiana multivariante, es decir,

$$q(\mathbf{x}) = \mathcal{N}(\mathbf{x} | \langle \mathbf{x} \rangle, \Sigma_{\mathbf{x}}), \quad (4.41)$$

donde su media y matriz de covarianza vienen dadas por

$$\langle \mathbf{x} \rangle = \Sigma_{\mathbf{x}} \left(\lambda_1 \langle \beta_1 \rangle \langle \mathbf{H} \rangle^T \mathbf{y}_1 + \langle \beta_2 \rangle \mathbf{y}_2 \right), \quad (4.42)$$

$$\Sigma_{\mathbf{x}}^{-1} = \langle \alpha \rangle (\Delta^h)^T \mathbf{W} (\Delta^h) + \langle \alpha \rangle (\Delta^v)^T \mathbf{W} (\Delta^v) + \lambda_1 \langle \beta_1 \rangle \langle \mathbf{H}^T \mathbf{H} \rangle + \langle \beta_2 \rangle \mathbf{I}, \quad (4.43)$$

con

$$w_j = \left\langle (\Delta_j^h(\langle \mathbf{x} \rangle))^2 + (\Delta_j^v(\langle \mathbf{x} \rangle))^2 \right\rangle, j = 1, \dots, N, \quad (4.44)$$

$$\mathbf{W} = \text{diag} \left(\frac{1}{\sqrt{w_j} + \epsilon} \right), j = 1, \dots, N, \quad (4.45)$$

con ϵ , un valor real positivo, suficientemente pequeño que evita la división por cero. La media $\langle \mathbf{x} \rangle$ de la distribución $q(\mathbf{x})$ va a coincidir con nuestra estimación de la imagen, y será calculada mediante el método de gradiente conjugado sobre la Ecuación (4.42). La matriz \mathbf{W} en (4.45) representa una matriz de adaptividad espacial que controla el nivel de suavizado en cada píxel dependiendo de la variación de intensidad en la posición del píxel, según indique el vector \mathbf{w} el cual representa la variación total de la imagen estimada y controla el equilibrio entre las fronteras de los objetos y el suavizado de la imagen estimada. Adicionalmente, el método controla

la aportación de cada una de las observaciones a la imagen estimada a través de los parámetros $\langle \beta_1 \rangle$ y $\langle \beta_2 \rangle$. Debemos mencionar que los parámetros son estimados simultáneamente con la imagen, como mostramos más abajo, y el proceso de estimación es guiado automáticamente hacia la imagen más probable. La mejor restauración se consigue mediante las restricciones impuestas en el modelo de emborronamiento y de imagen. Por ejemplo, si la observación de corta exposición es ruidosa en exceso, el método tenderá a confiar más en la observación borrosa, aumentando $\langle \beta_1 \rangle$ y decrementando $\langle \beta_2 \rangle$, y adaptando el parámetro $\langle \alpha \rangle$ para conservar las fronteras sin amplificar el ruido.

A continuación buscamos la aproximación de las distribuciones sobre los coeficientes de la PSF $q(h_j)$ a partir de (4.40), $q(h_j)$ forma una distribución Gaussiana rectificada, dada por

$$q(h_j) = \mathcal{N}^R(h_j | \hat{h}_j, \tilde{h}_j), \quad (4.46)$$

de parámetros

$$\hat{h}_j = (\tilde{h}_j)^{-1} \left[- \sum_{d=1}^D \langle \sigma_d \rangle \mu_{jd} \lambda_1 \langle \beta_1 \rangle \sum_{n=1}^N \left\langle X_{nj} \left((y_1)_n - \sum_{\substack{m=1 \\ m \neq j}}^M X_{nm} h_m \right) \right\rangle \right. \\ \left. + (1 - \lambda_1) \langle \beta_{12} \rangle \sum_{n=1}^N (Y_2)_{nj} \left((y_1)_n - \sum_{\substack{m=1 \\ m \neq j}}^M (Y_2)_{nm} \langle h_m \rangle \right) \right], \quad (4.47)$$

$$\tilde{h}_j = \lambda_1 \langle \beta_1 \rangle \sum_{n=1}^N \langle X_{nj}^2 \rangle + (1 - \lambda_1) \langle \beta_{12} \rangle \sum_{n=1}^N (Y_2)_{nj}^2, \quad (4.48)$$

donde \mathbf{X} e \mathbf{Y}_2 son matrices de convolución construidas a partir de \mathbf{x} e \mathbf{y}_2 , respectivamente, $(\cdot)_{ij}$ denota el elemento (i, j) -ésimo de una matriz, y hemos tenido en cuenta que $\lambda_2 = 1 - \lambda_1$. La esperanza $\langle h_j \rangle$ de las distribuciones $q(h_j)$, es decir, nuestra estimación puntual para h_j , viene dada por [51]

$$\langle h_j \rangle = \hat{h}_j + \sqrt{\frac{2}{\pi \tilde{h}_j}} \frac{1}{\operatorname{erfcx}(-\hat{h}_j \sqrt{\frac{\tilde{h}_j}{2}})}, \quad (4.49)$$

$$\operatorname{var}(h_j) = \tilde{h}_j^{-1} + \sqrt{\frac{1}{\pi \tilde{h}_j}} \frac{\hat{h}_j}{\operatorname{erfcx}(-\hat{h}_j \sqrt{\frac{\tilde{h}_j}{2}})}, \quad (4.50)$$

donde $\operatorname{erfcx}(\cdot)$ representa la función de error complementario escalado.

A continuación tenemos que calcular las distribuciones de los hiperparámetros $\alpha, \beta_1, \beta_2, \beta_{12}, \sigma_d$

en (4.38), (4.39), y (4.40) como

$$q(\alpha) = \text{Gamma}(\alpha | \bar{a}_\alpha, \bar{b}_\alpha),, \quad (4.51)$$

$$q(\beta_1) = \text{Gamma}(\beta_1 | \bar{a}_{\beta_1}, \bar{b}_{\beta_1}), \quad (4.52)$$

$$q(\beta_2) = \text{Gamma}(\beta_2 | \bar{a}_{\beta_2}, \bar{b}_{\beta_2}), \quad (4.53)$$

$$q(\beta_{12}) = \text{Gamma}(\beta_{12} | \bar{a}_{\beta_{12}}, \bar{b}_{\beta_{12}}). \quad (4.54)$$

$$q(\sigma_d) = \text{Gamma}(\sigma_d | \bar{a}_{\sigma_d}, \bar{b}_{\sigma_d}), \quad (4.55)$$

Obsérvese que la aproximación a la distribución a posteriori tiene la misma forma que la correspondiente distribución a priori debido al uso de distribuciones conjugadas. Utilizamos las medias de estas distribuciones como estimadores, lo que conduce a

$$\langle \alpha \rangle = \frac{\bar{a}^{(\alpha)}}{\bar{b}^{(\alpha)}} = \frac{a^{(\alpha)} + \frac{N}{2}}{b^{(\alpha)} + \sum_j \sqrt{w_j}}, \quad (4.56)$$

$$\langle \beta_1 \rangle = \frac{\bar{a}^{(\beta_1)}}{\bar{b}^{(\beta_1)}} = \frac{a^{(\beta_1)} + \frac{N}{2}}{b^{(\beta_1)} + \frac{1}{2} \langle \|\mathbf{y}_1 - \mathbf{H}\mathbf{x}\|^2 \rangle}, \quad (4.57)$$

$$\langle \beta_2 \rangle = \frac{\bar{a}^{(\beta_2)}}{\bar{b}^{(\beta_2)}} = \frac{a^{(\beta_2)} + \frac{N}{2}}{b^{(\beta_2)} + \frac{1}{2} \langle \|\mathbf{y}_2 - \mathbf{x}\|^2 \rangle}, \quad (4.58)$$

$$\langle \beta_{12} \rangle = \frac{\bar{a}^{(\beta_{12})}}{\bar{b}^{(\beta_{12})}} = \frac{a^{(\beta_{12})} + \frac{N}{2}}{b^{(\beta_{12})} + \frac{1}{2} \langle \|\mathbf{y}_1 - \mathbf{Y}_2 \mathbf{h}\|^2 \rangle}, \quad (4.59)$$

$$\langle \sigma_d \rangle = \frac{\bar{a}^{(\sigma_d)}}{\bar{b}^{(\sigma_d)}} = \frac{a^{(\sigma_d)} + \sum_j \mu_{jd}}{b^{(\sigma_d)} + \sum_j \mu_{jd} \langle h_j \rangle}. \quad (4.60)$$

Además

$$q(\{\tau_d\}_{d=1}^D) = \text{Dirichlet}(\{\tau_d\}_{d=1}^D | \{\bar{c}_{\tau_d}\}_{d=1}^D), \quad (4.61)$$

donde $\bar{c}_{\tau_d} = c_{\tau_d}^o + \sum_j \mu_{jd}$ y entonces

$$\langle \tau_d \rangle = \frac{\bar{c}_{\tau_d}}{\sum_{d=1}^D \bar{c}_{\tau_d}}. \quad (4.62)$$

Finalmente, las variables auxiliares, μ_{jd} se calculan de nuevo a partir de (4.40) como

$$\mu_{jd} \propto \langle \tau_d \rangle \text{Expon}(\langle h_j \rangle | \langle \sigma_d \rangle), \quad j = 1, \dots, M \quad (4.63)$$

sujeto a la condición

$$\sum_{d=1}^D \mu_{jd} = 1, \quad j = 1, \dots, M \quad (4.64)$$

Algoritmo 1: Algoritmo de deconvolución para emborronamiento invariante

Input: Imagen con tiempo correcto \mathbf{y}_1 , y de exposición corta \mathbf{y}_2 .

Output: Estimación de la imagen \mathbf{x} , emborronamiento \mathbf{h} y los parámetros necesarios

Establecer la estimación inicial de la imagen $\langle \mathbf{x} \rangle^{(0)} = \mathbf{y}_1$.

Calcular las estimaciones iniciales de $\langle h_j \rangle$, β_1 , β_2 , β_{12} , α , σ_d and τ_d usando $\langle \mathbf{x} \rangle^{(0)}$, \mathbf{y}_1 , \mathbf{y}_2 y λ .

for $k = 1, 2, \dots$ hasta convergencia **do**

 Encontrar la distribución de la imagen $q^k(\mathbf{x})$ usando (4.42) y (4.43).

 Encontrar coeficientes de las distribuciones de la PSF $q^k(h_j)$ usando (4.47)-(4.50).

 Encontrar la estimación de los hiperparámetros de las distribuciones (4.51)-(4.55) y (4.61).

 Encontrar las variables auxiliares $\{\mu_{jd}\}$ usando (4.63).

end

El cálculo explícito de las esperanzas $\langle \mathbf{H}^t \mathbf{H} \rangle$ en 4.43, $\langle (\Delta_i^u(\mathbf{x}))^2 + (\Delta_i^v(\mathbf{x}))^2 \rangle$ en 4.44, $\langle \|\mathbf{y}_1 - \mathbf{H}\mathbf{x}\|^2 \rangle$ en 4.57, $\langle \|\mathbf{y}_2 - \mathbf{x}\|^2 \rangle$ en 4.58 y $\langle \|\mathbf{y}_1 - \mathbf{Y}_2 \mathbf{h}\|^2 \rangle$ en 4.59, viene dado por

$$\langle \mathbf{H}^t \mathbf{H} \rangle = \langle \mathbf{H}^t \rangle \langle \mathbf{H} \rangle + \Upsilon_{\mathbf{h}}, \quad (4.65)$$

$$\left\langle \left(\Delta_i^h(\mathbf{x}) \right)^2 + \left(\Delta_i^v(\mathbf{x}) \right)^2 \right\rangle = \left(\Delta_i^h(\langle \mathbf{x} \rangle) \right)^2 + \left(\Delta_i^v(\langle \mathbf{x} \rangle) \right)^2 + \text{trace} \left(\left((\Delta^h)^t \Delta^h + (\Delta^v)^t \Delta^v \right) \Sigma_{\mathbf{x}} \right), \quad (4.66)$$

$$\langle \|\mathbf{y}_1 - \mathbf{H}\mathbf{x}\|^2 \rangle = \|\mathbf{y}_1 - \langle \mathbf{H} \rangle \langle \mathbf{x} \rangle\|^2 + \text{trace}(\langle \mathbf{H} \rangle^t \langle \mathbf{H} \rangle \Sigma_{\mathbf{x}}) + \text{trace}(\langle \mathbf{x} \rangle^t \langle \mathbf{x} \rangle \Upsilon_{\mathbf{h}}) + \text{trace}(\Sigma_{\mathbf{x}} \Upsilon_{\mathbf{h}}), \quad (4.67)$$

$$\langle \|\mathbf{y}_2 - \mathbf{x}\|^2 \rangle = \|\mathbf{y}_2 - \langle \mathbf{x} \rangle\|^2 + \text{trace}(\Sigma_{\mathbf{x}}), \quad (4.68)$$

$$\langle \|\mathbf{y}_1 - \mathbf{Y}_2 \mathbf{h}\|^2 \rangle = \|\mathbf{y}_1 - \mathbf{Y}_2 \langle \mathbf{h} \rangle\|^2 + \text{trace}(\mathbf{Y}_2^t \mathbf{Y}_2 \Upsilon_{\mathbf{h}}). \quad (4.69)$$

donde

$$\Upsilon_{\mathbf{h}} = \text{diag} \left(\sum_{j=1}^M \text{var}(h_j) \right), \quad (4.70)$$

con $\text{var}(h_j)$ expresado en la Ecuación (4.50).

El método que proponemos se detalla en el algoritmo 1. En relación a la implementación del algoritmo es importante hacer notar que la matriz de covarianzas $\Sigma_{\mathbf{x}}$ únicamente es necesaria calcularla de forma explícita en las Ecs. (4.66), (4.67) y (4.68); debido a las dimensiones de esta matriz ($N \times N$) su cálculo no es viable y por tanto $\Sigma_{\mathbf{x}}^{-1}$ será aproximada por una matriz diagonal cuyos elementos de $\Sigma_{\mathbf{x}}$ se calculan por el inverso de los elementos de la diagonal principal, especificados en 4.43. Se han llevado a cabo diversos experimentos en imágenes de pequeño tamaño en los que calculamos explícitamente $\Sigma_{\mathbf{x}}$ y comparamos con la aproximación que proponemos,

llegando a la conclusión de que produce resultados muy parecidos y presenta un mínimo efecto en el proceso de estimación. Téngase en cuenta que otros métodos bayesianos han aplicado aproximaciones similares [77, 28, 78]. Nótese que en la Ecuación 4.42, utilizamos el método de gradiente conjugado y por tanto, no es necesario el cálculo explícito de la matriz de covarianza.

La estrategia que se ha seguido en este trabajo estima las distribuciones de las variables desconocidas en lugar de estimadores puntuales; la ventaja más clara es que nos permite incorporar la incertidumbre al proceso de estimación a través de las esperanzas en (4.65)-(4.69), de forma que para calcular cada una de las incógnitas tenemos en cuenta los errores que puedan producirse en el resto de variables. Por otra lado, se podrían haber simplificado los cálculos y la complejidad del algoritmo asumiendo distribuciones degeneradas, es decir distribuciones en las que un único valor toma el valor máximo y el resto se deja a cero (función delta), haciendo coincidir el delta con la moda de la distribución. En este caso las covarianzas valdrían cero y equivale al estimador MAP de las incógnitas.

4.6. Experimentos

El método propuesto ha sido evaluado tanto en imágenes degradadas sintéticamente como en imágenes reales. En un primer experimento, las imágenes sintéticas se utilizan para probar la exactitud de la estimación tanto de forma numérica como visual y demostrar que la combinación de modelos de observación propuesto proporciona mejores resultados que con un único modelo o incluso asumiendo que los modelos son independientes como en [2]. Los resultados obtenidos con el algoritmo propuesto se comparan con el método de Richardson-Lucy [29, 30] y con el método reciente de exposición dual en [2]. Adicionalmente comparamos con el estado del arte en métodos de supresión de ruido en [1] aplicado a la observación ruidosa.

En todos los experimentos se utilizaron los siguientes valores para la inicialización de los parámetros. Como ya fue descrito en el Alg. 1, establecemos la imagen observada \mathbf{y}_1 como la estimación inicial de \mathbf{x} , $\langle \mathbf{x} \rangle^0$. Elegimos usar dos clases para el emborronamiento, es decir $D = 2$, lo que significa que el emborronamiento estará formado por píxeles con dos tipos de valores, unos próximos a cero y otros con valores altos. Los parámetros de forma e inversa de la escala para la distribuciones Gamma se fijan a un valor pequeño (0,001), y $c_{\tau_d}^0$ se fija a 1, consiguiendo con ello que el proceso de estimación se base más en las observaciones que en el conocimiento previo que podamos tener sobre los parámetros.

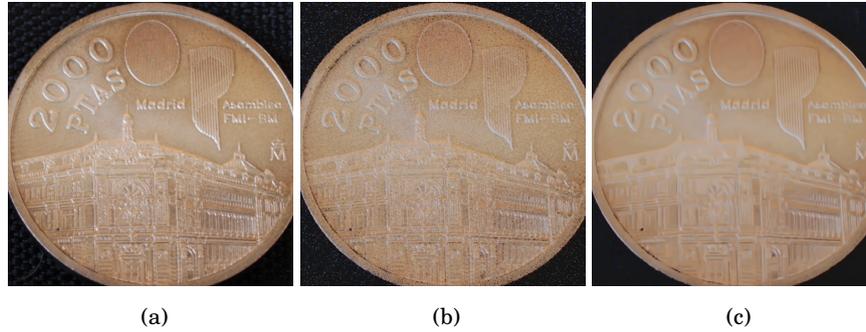


Figura 4.2: (a) Imagen original. (b) Observación ruidosa simulando imagen de corta exposición tras la calibración. (c) Método de supresión de ruido en [1] aplicado sobre la imagen ruidosa.

La estimación inicial de la función de emborronamiento se obtiene como $\text{argmin}_{\mathbf{h}} \|\mathbf{Y}_2 \mathbf{h} - \mathbf{y}_1\|^2$, que puede calcularse de forma eficiente en el dominio de Fourier. El tamaño del emborronamiento M se elige como el menor número de píxeles que cubre las entradas más importantes de la estimación inicial de la PSF. El valor inicial de los parámetros β_1 , β_2 , β_{12} y α se calculan a partir de las Ecuaciones 4.56 - 4.59. Como ya hemos comentado previamente, suponemos dos clases distintas para los elementos de la PSF; por lo tanto el valor inicial de τ_d lo fijamos a la proporción de píxeles de la estimación de la PSF inicial que son más pequeños o mayores que la mitad del valor máximo de \mathbf{h} , para $d = 1$ y $d = 2$, respectivamente, con σ_d la inversa de la media de los píxeles de cada clase. Finalmente, $\{\mu_{jd}\}$ se calcula utilizando (4.63). En los experimentos se ha variado el parámetro λ_1 de 0 a 1 con una precisión de 0,1 ejecutando el Algoritmo 1 hasta cumplir el criterio de convergencia $\|\langle \mathbf{x} \rangle^i - \langle \mathbf{x} \rangle^{i-1}\|^2 / \|\langle \mathbf{x} \rangle^{i-1}\|^2 < 10^{-5}$.

Queremos hacer notar que exclusivamente restauramos la banda de luminancia de las imágenes en color. La imagen a color resultante se compone de la banda de luminancia restaurada junto a las bandas de crominancia de la imagen borrosa original.

4.6.1. Experimentos sintéticos

Para los experimentos sintéticos, generamos el par de imágenes a partir de la imagen original que se muestra en la Figura 4.2a. Con el fin de simular una imagen subexpuesta realista, seguimos el diagrama de adquisición de imágenes descrito en la Sección 1.2 y simulamos la corrupción por ruido, de la imagen subexpuesta, en el dominio de irradiancia. En primer lugar, la imagen original en la Figura 4.2a se convirtió al dominio de irradiancia deshaciendo la correc-

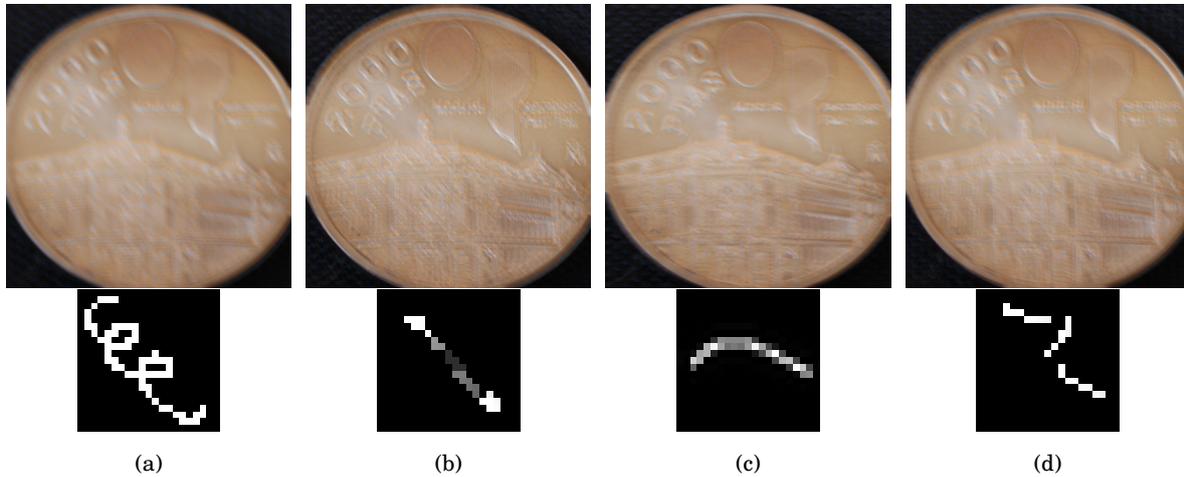


Figura 4.3: Imágenes borrosas simulando la imagen de exposición correcta en entornos de baja luminosidad. La PSF utilizada para generar las observaciones se muestra debajo de cada una de las imágenes. Nótese que los valores de las funciones de emborronamiento se han llevado al rango $[0,255]$ por razones de visualización.

ción gamma, suponiendo $\gamma = 2,2$. Posteriormente, asumiendo que la imagen de corta exposición posee un valor de exposición (EV) seis pasos inferior a la imagen original, dividimos cada valor de pixel por $2^6 = 64$. Añadimos ruido gaussiano de varianza 199,3, correspondiente a una SNR de 10 dB. Finalmente, a la imagen resultante se le aplicó una corrección gamma con $\gamma = 2,2$ y se cuantificó a ocho bits para obtener la imagen subexpuesta simulada de la Figura 4.2b.

Las imágenes borrosas, mostradas en la Figura 4.3, se obtuvieron mediante la convolución de la imagen original con varias PSFs que aparecen en la segunda fila de la Figura 4.3, y la adición de ruido gaussiano de media cero y varianza 0,35 para obtener una SNR de 40dB. El par de imágenes formado por la imagen borrosa y la ruidosa son las entradas del algoritmo.

Utilizamos la métrica PSNR (Peak Signal-to-Noise Ratio o relación señal-ruido pico) para evaluar numéricamente el funcionamiento del algoritmo. La PSNR se expresa en escala logarítmica, utilizando como unidad el decibelio (dB) y se define como

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}_y^2}{\text{MSE}} \right) = 20 \log_{10} \left(\frac{\text{MAX}_y}{\sqrt{\text{MSE}}} \right), \quad (4.71)$$

con el error cuadrático medio o MSE (Mean Square Error) definido como

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N \| \mathbf{y} - \mathbf{x} \|^2, \quad (4.72)$$

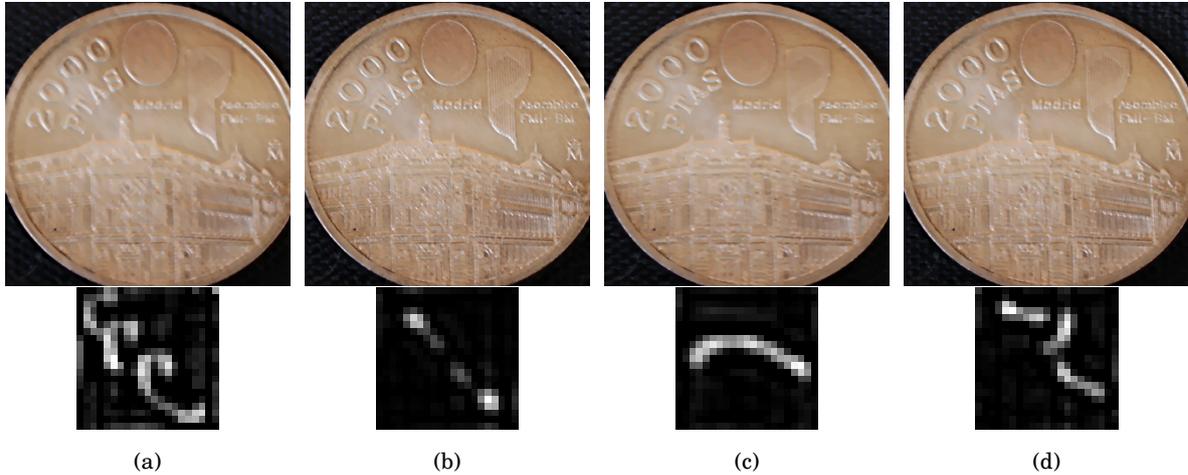


Figura 4.4: Resultados de aplicar nuestro método seleccionando el λ óptimo sobre las imágenes borrosas de la Figura 4.3 y ruidosa 4.2b, debajo se muestra la PSF estimada en cada experimento.

donde N representa el número de píxeles de la banda de luminancia de las imágenes \mathbf{y} y \mathbf{x} . MAX_y denota el valor máximo que puede tomar un píxel en la imagen y que en nuestro caso por tratarse de imágenes de 8 bit de profundidad, sería de $2^8 - 1 = 255$. Una mayor PSNR denota mejor calidad de imagen.

El gráfico de la Figura 4.5 muestra la evolución de la PSNR en los 4 experimentos sintéticos realizados con el algoritmo propuesto, en el que variamos λ_1 de 0 a 1, con una precisión de 0,1. Los valores de PSNR se pueden consultar en la Tabla 4.1, mientras que las restauraciones con el λ óptimo (en términos de una mayor PSNR) se presentan en la Figura 4.4. En los experimentos sintéticos realizados, se obtiene una mejor restauración con valores de λ_1 comprendidos entre 0,5 y 0,7 excepto en el resultado del primer experimento, en la Figura 4.4a donde la mayor PSNR se obtiene usando $\lambda_1 = 0,9$. Esto puede explicarse debido a la peor estimación de la PSF en ese experimento concreto. Como se puede observar, valores de λ pequeños producen restauraciones suaves mientras que cuando aumenta el valor de λ las restauraciones presentan bordes más realzados pero también más ruidosas.

Cabe destacar que el mejor resultado se obtuvo mediante una combinación no degenerada de las divergencias, es decir, la combinación de ambos modelos proporciona mejores resultados que la utilización de un único modelo. Téngase en cuenta que $\lambda_1 = \lambda_2 = 0,5$ no es equivalente al modelo en [2], que corresponde a $\lambda_1 = \lambda_2 = 1$. Con el fin de comparar ambos métodos, ejecuta-

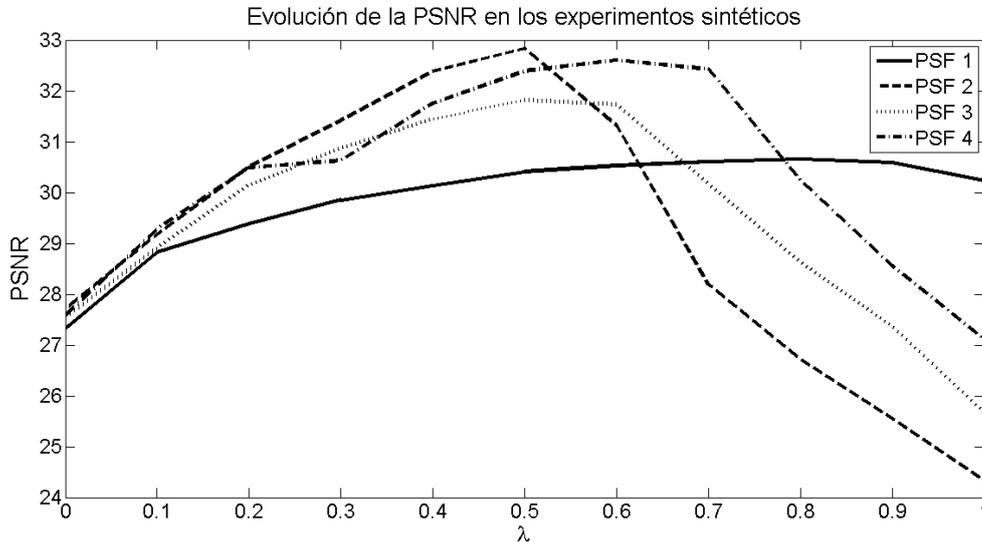


Figura 4.5: Evolución de la PSNR en función de λ_1 para los resultados de la Figura 4.4.

mos los mismos experimentos con el algoritmo en [2], cuya restauraciones y PSFs se muestran en la Figura 4.6 obteniéndose las PSNRs que se muestran en la Tabla 4.1. El algoritmo propuesto proporciona siempre los mejores resultados en PSNR. En cuanto a calidad visual, el algoritmo propuesto proporciona resultados algo más suaves que el algoritmo en [2], debido a una estimación menos ruidosa (pero algo más suave) de la PSF, que la obtenida por el método [2]. Respecto al resultado de aplicar el algoritmo de RL, implementado en MATLAB mediante la rutina `deconvblind`, podemos apreciar claramente un exceso de efecto de anillado (Fig. 4.7), incluso inicializando el método RL con nuestra estimación del núcleo. El número de iteraciones para la ejecución de `deconvblind` fue fijado a 20. Nótese también que las estimaciones de la PSF proporcionadas por `deconvblind` no han mejorado las PSFs iniciales introducidas, que son las proporcionadas por nuestro método, lo que indica que la estimación producida por el método propuesto es suficientemente buena.

4.6.2. Experimentos reales

En la siguiente serie de experimentos, probamos nuestro algoritmo en situaciones reales. Para ello hemos realizado dos tomas con diferentes tiempos de exposición, en entornos de baja luminosidad usando varias cámaras digitales distintas. Los pares de imágenes se muestran en las Figuras 4.8a y 4.8b, 4.9a y 4.9b. Las imágenes de entrada han sido calibradas fotométricamente.

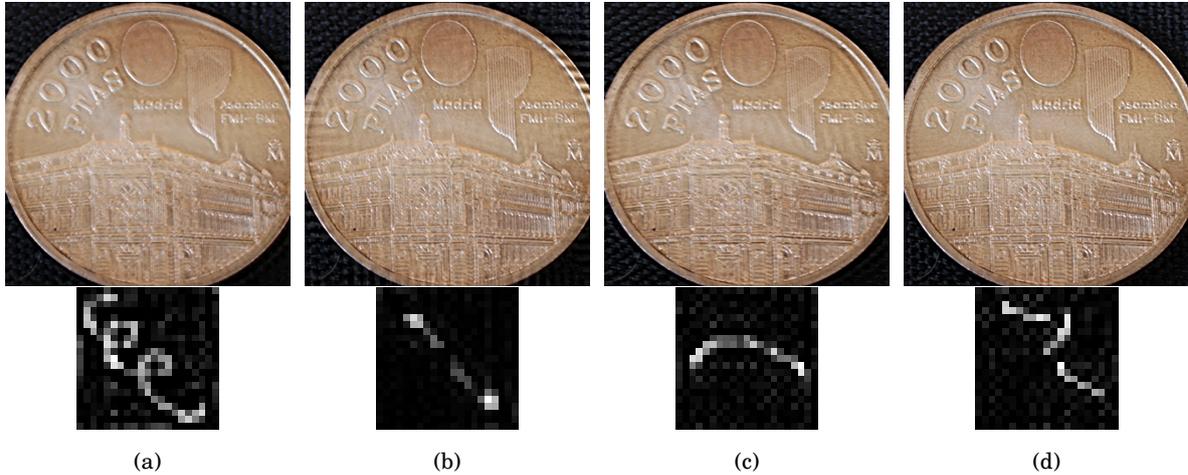


Figura 4.6: Resultados de aplicar el método [2] a las imágenes borrosas de la Figura 4.3 y 4.2b, debajo se muestra la PSF estimada.

camente y geoméricamente tal y como se describe en el Capítulo 3. Nótese que, en nuestro algoritmo, la ecualización de histograma se realiza exclusivamente en la banda de luminancia; sin embargo, para mostrar las Figuras 4.8b y 4.9b, las bandas de crominancia de la imagen de corta exposición fueron también calibradas fotométricamente usando la técnica de ecualización de histograma. En todos los experimentos se utilizó umbral de convergencia en iteraciones consecutivas de $threshold = 10^{-5}$. Nótese que los experimentos reales han sido comparados con el estado del arte en exposición dual cuyos detalles se encuentran en [2].

El primer experimento se llevó a cabo sobre el par de imágenes de las Figuras 4.8a y 4.8b, de tamaño 512×512 . Las imágenes fueron captadas con la cámara Nikon D5000. El experimento exhibe una escena con emborronamiento por movimiento de cámara. Para captar la imagen ruidosa de baja exposición se seleccionó un tiempo de $1/200$ segundos y 400 de valor ISO, mientras que el tiempo de exposición para la imagen borrosa de exposición correcta fue de $1/3$ segundos e ISO 200. Ambos disparos con una apertura fija de $f/8$. Se seleccionó un tamaño de PSF de 21×21 píxeles para la ejecución del algoritmo en [2] y el propuesto. Las Figuras 4.8c y 4.8d presentan las restauraciones y PSFs estimadas por el método en [2] y el propuesto, respectivamente. A la vista de los resultados, la imagen restaurada por el método en [2] aparenta mayor contraste pero sólo consiguió eliminar el emborronamiento parcialmente; mientras que el método propuesto proporciona detalles más nítidos y preserva las texturas, eliminando la mayor parte del

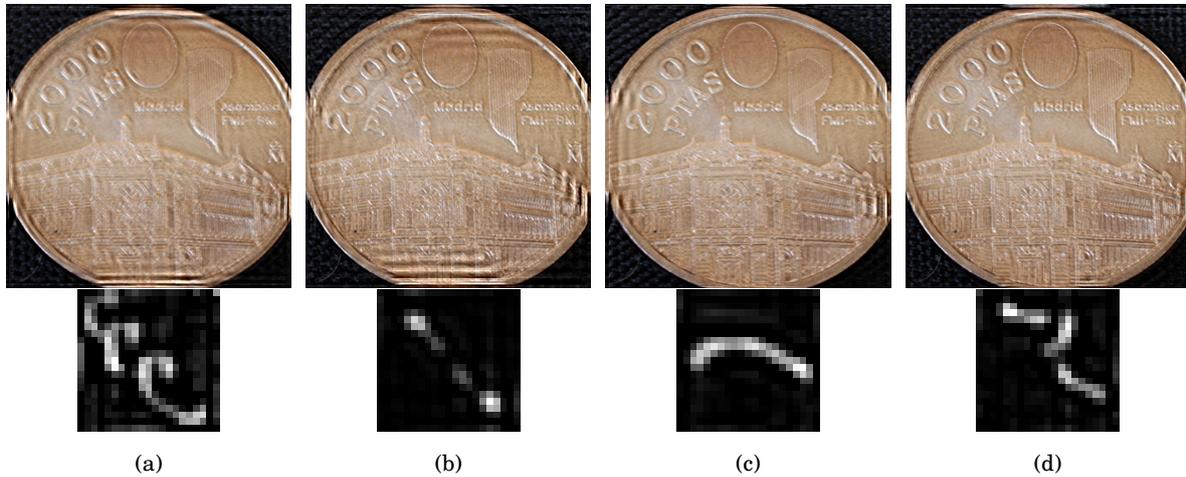


Figura 4.7: Resultados de aplicar el método `deconvblind` usando la PSF estimada por el método propuesto a las imágenes borrosas de la Figura 4.3 y 4.2b.

emborronamiento. Véase, por ejemplo como los títulos en el tomo de los libros son legibles.

En el segundo experimento se utilizaron el par de imágenes de las Figuras 4.9a y 4.9b de tamaño 800×600 . La imagen de exposición correcta presenta dos tipos de emborronamientos: el primero de ellos debido al movimiento de la cámara y un segundo emborronamiento producido por la reducida profundidad de campo. Las imágenes se tomaron con una cámara Canon EOS 500D. El tiempo de exposición de la imagen borrosa es $1/10$ s mientras que en la imagen de menor exposición se utilizó $1/160$ s con una compensación de exposición (EV) de -4, el factor ISO en ambas tomas se fijó a 100. La imagen presenta un emborronamiento considerable, por ello se utilizó un tamaño de PSF de 31×31 píxeles. La imagen de menor exposición de la Figura 4.9b contiene ruido de cuantificación que deseamos evitar en la restauración. Desafortunadamente, ninguno de los métodos parece efectivo frente a este tipo de emborronamiento (véase Figuras 4.9c y 4.9d). Además únicamente fue posible obtener una estimación de la PSF apta para resolver el emborronamiento en la parte del tejado de la casa al fondo de la imagen pero parcialmente. Esto es debido a que el emborronamiento varía según la zona de la imagen. Este hecho fue el desencadenante para proponer los métodos del capítulo siguiente, que nos van a permitir modelar el emborronamiento espacialmente variante permitiendo una mayor flexibilidad.

Cuadro 4.1: Métrica PSNR (Peak Signal to Noise Ratio) para las imágenes de los experimento sintéticos

	Conjunto de PSFs			
	a	b	c	d
Borrosa	27.17	26.58	26.87	26.41
Ruidosa	26.60			
Denoising [1]	29.30			
deconvblind	23.74	23.20	26.46	26
Método en [2]	30.64	29.18	28.37	29.15
Propuesto (λ óptimo)	30.65	32.83	31.82	32.60

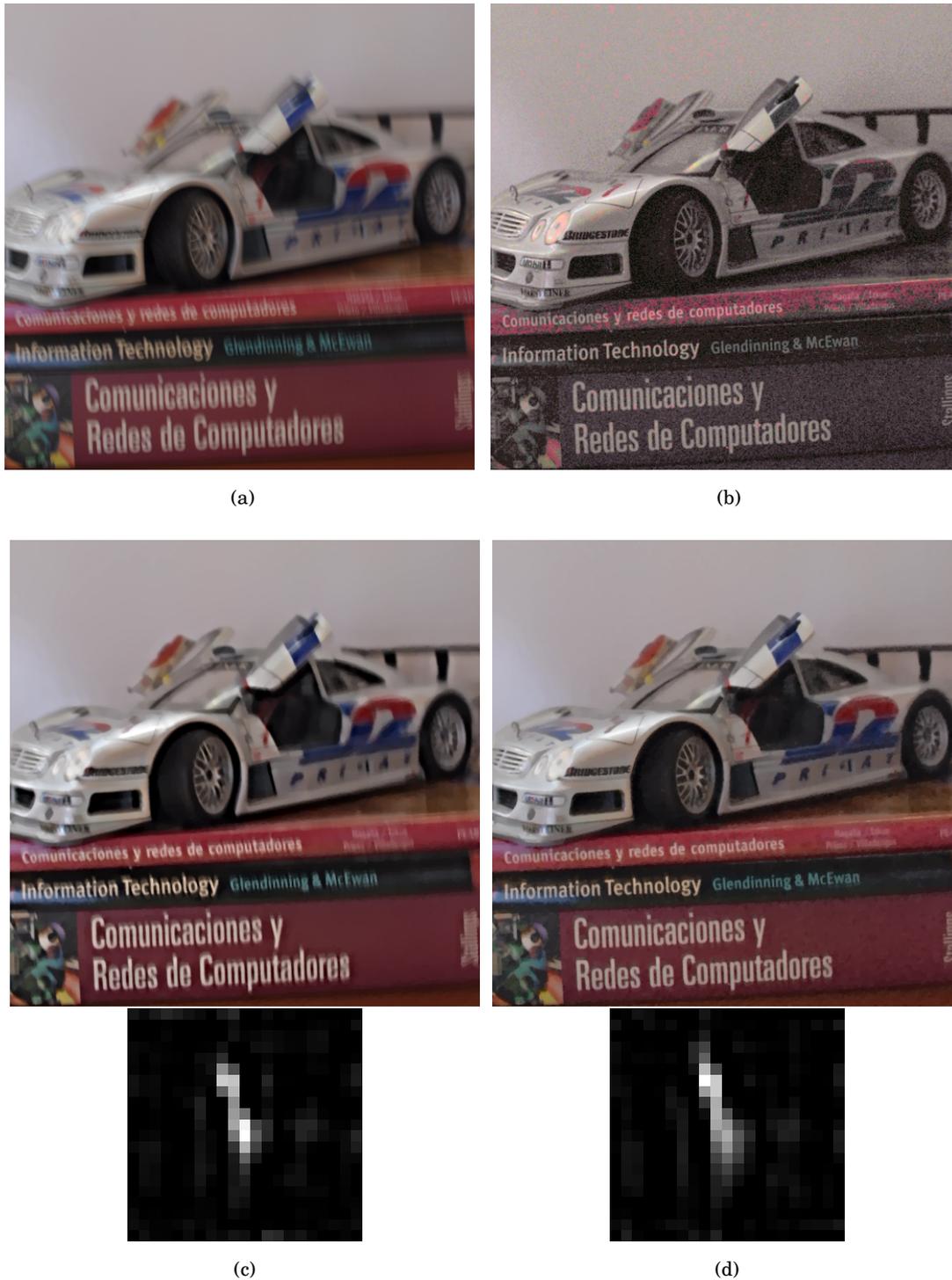


Figura 4.8: Experimento real 1: (a) observación borrosa, (b) ruidosa después de la calibración (c) restauración y PSF obtenidas por la aplicación del método en [2], (d) restauración y PSF obtenidas utilizando el método propuesto con $\lambda_1 = 0,8$.



Figura 4.9: Experimento real 2: (a) observación borrosa, (b) ruidosa después de la calibración (c) restauración y PSF obtenidas por la aplicación del método en [2], (d) restauración y PSF obtenidas utilizando el método propuesto con $\lambda_1 = 0,8$.

Capítulo 5

Métodos de deconvolución basados en núcleo espacialmente variante

En los momentos de crisis, sólo la imaginación es más importante que el conocimiento.

Albert Einstein (1879-1955).

Físico, considerado el científico más importante del siglo XX.

5.1. Introducción

Como ya hemos visto en el capítulo anterior la eliminación de emborronamiento de la imagen puede ser una tarea bastante compleja, en la que utilizar una segunda observación además de la imagen borrosa ayuda a obtener un emborronamiento más preciso y por consiguiente, una restauración más nítida. Sin embargo también añade otros problemas como la necesidad de calibración de las imágenes y el ruido. En dicho capítulo se propuso una solución efectiva para eliminar el emborronamiento por movimiento de cámara pero, como se mostró, insuficiente en casos de emborronamiento espacialmente variante. Como se explica en [39] el emborronamiento creado por cámaras de fotografía convencionales rara vez es espacialmente invariante, ya que el sistema óptico de la cámara genera un considerable emborronamiento espacialmente variante más acusado en los bordes de la imagen [39]. Por ello, proponemos en el presente capítulo un marco para la eliminación de un tipo de emborronamiento más realista, que varía

espacialmente.

Los métodos propuestos en este capítulo procesan la imagen por bloques de tamaño reducido, en los que el emborronamiento se considera espacialmente invariante pero, al considerar la imagen completa, es globalmente variante. Para ello, primero se dividen las dos imágenes de entrada, borrosa y ruidosa, en bloques solapados, calculando una función de emborronamiento en cada uno de los pares de bloques. El solapamiento entre bloques permite una mayor variabilidad especialmente en las zonas solapadas. Una vez estimada la PSF en cada bloque, se deconvoluciona el bloque borroso, combinando la información tanto del bloque borroso como del ruidoso, obteniendo una restauración que habrá que fusionar posteriormente con el resto de restauraciones de los bloques para obtener una imagen restaurada.

En el presente capítulo se proponen dos métodos de restauración basados en la técnica descrita y que además presentan la capacidad de elegir entre la restauración obtenida de la deconvolución y la obtenida tras eliminación de ruido únicamente sobre la imagen ruidosa. Esto va a ser muy útil en bloques donde la función de emborronamiento no pudo ser estimada correctamente. Otra ventaja de este enfoque de deconvolución con emborronamiento espacialmente variante, reside en una mayor flexibilidad, ya que el uso de bloques solapados permite modelar emborronamientos más complejos, de forma que aunque la PSF estimada en cada bloque va a tratarse de forma invariante, las regiones solapadas van a tener una mezcla de las restauraciones calculadas en bloques adyacentes.

Las diferencias de los métodos propuestos en este capítulo con el algoritmo de emborronamiento espacialmente invariante que presentamos en el Capítulo 4 son varias: La primera y más substancial reside en el modelo de emborronamiento, que en este caso va a permitir un modelo más realista que varía espacialmente; pero también en la forma de estimarlo, ya que no se reestima en cada iteración sino que partimos de una única estimación al inicio del procedimiento. Otra diferencia importante que los métodos propuestos en este capítulo contienen etapas de corrección del núcleo o detección de deconvoluciones pobres que los hacen robustos frente a estimaciones poco fiables. La división en bloques con solapamiento también presenta una serie de dificultades a las que damos solución a lo largo del capítulo. Por último, contamos con la aportación del método de supresión de ruido en zonas en las que no fue posible una estimación correcta del emborronamiento. Proponemos dos alternativas para la estrategia de selección de la restauración final, una para cada uno de los algoritmos propuestos en el presente capítulo.

Algoritmo 2: Algoritmo general

Pre-procesamiento: Calibración del par de imágenes y división en bloques (Sec. 5.2)

Estimación del núcleo de emborronamiento (Sec. 5.3)

Deconvolución/supresión de ruido del bloque. (Secciones 5.4 y 5.5)

Pos-procesamiento: Fusionar los bloques restaurados para conformar la imagen final. (Sec. 5.6)

La organización del capítulo sigue los pasos descritos en el algoritmo general mostrado en el Algoritmo 2. En la sección 5.2 describimos la fase de preprocesamiento. Posteriormente en la sección 5.3 proponemos el método de estimación de emborronamiento, compartido por los dos algoritmos de restauración presentados en este capítulo y aplicado a cada uno de los pares de bloques. La sección 5.4.2 detalla el proceso de deconvolución rápido con reparación de PSFs, que ha sido publicado en la 19th European Signal Processing Conference 2011 (EUSIPCO) [7]. Alternativamente se ha desarrollado otro método de deconvolución que explicamos en la sección 5.5 y que ha sido publicado recientemente en la revista *Information Fusion* [9]. Ambos métodos hacen uso de un marco Bayesiano jerárquico y aplican inferencia variacional para encontrar una aproximación a la distribución a posteriori en cada una de las parejas de bloques. En el primer método, se realiza un procedimiento de reparación de PSF previo a la aplicación del método de deconvolución rápido, basado en un modelo de imagen TV y distribuciones propias sobre los parámetros. En el segundo método de deconvolución se propone una combinación de modelos a priori de imagen, en un enfoque similar al descrito en el capítulo anterior, en el que, además, la aproximación de la distribución a posteriori de la imagen se lleva a cabo mediante la técnica de gradiente conjugado evitando el cálculo explícito de la matriz de covarianza pero haciéndolo más lento que el método de deconvolución rápido. La última de las etapas del algoritmo principal fusiona los bloques y se detalla en la sección 5.6. Para concluir el presente capítulo mostramos la sección de experimentos en la sección 5.7.

5.2. Preprocesamiento

Como ya se explicó ampliamente en el capítulo 3, el par de imágenes con diferente exposición se deben calibrar fotométrica y geoméricamente. Para la calibración fotométrica, se aplica ecualización de histograma [63] en la imagen de corta exposición apoyándonos en el histograma de la imagen de larga exposición. Esta técnica de ecualización del histograma es adecuada para

nuestro propósito, ya que trabaja con los histogramas acumulados, dejando de lado la geometría de la escena.

Una vez que ambas imágenes se encuentran calibradas fotométricamente, procedemos con la calibración geométrica. Esto se lleva a cabo mediante la extracción de los puntos en correspondencias de ambas imágenes con Surf [67] y la supresión de *outliers* (valor fuera de la serie) con RANSAC [68], para calcular la matriz de homografía con el algoritmo DLT [61] y finalmente alinear las imágenes aplicando la matriz de homografía previamente calculada. Llegado a este punto, es importante hacer notar que la precisión en el calibrado influye directamente en la restauración. Una mala calibración podría dar lugar a una estimación más compleja del núcleo de emborronamiento. Nótese que las imágenes están representadas en el espacio de color YC_bC_r . El método de restauración se aplica exclusivamente a la banda de luminancia. Finalmente la luminancia restaurada se fusiona con las bandas de crominancia de la imagen borrosa.

Una vez calibradas las imágenes de entrada, éstas se dividen en bloques solapados. De esta forma podremos modelar un emborronamiento espacialmente variante, mediante el cálculo de una única PSF para cada pareja de bloques. Este enfoque nos permite modelar un grupo más amplio de emborronamientos ya que las PSFs en las zonas solapadas van a estar formadas por la combinación de PSFs de los bloques aledaños. Nótese que el marco de bloques solapados podría utilizarse en otras áreas aparte de restauración de imágenes, tales como, eliminación de ruido o, recientemente, en superresolución para aumentar la resolución y eliminar ruido del mapa de profundidad de baja resolución que actualmente obtienen las cámaras comerciales como Kinect [79] o similares, combinando la información de un sensor convencional de color de alta resolución (véase el trabajo en [8]). Este marco también nos ayuda a optimizar la asignación de memoria y tiempo de cálculo, ya que utilizamos bloques de tamaño reducido. Aunque está fuera del ámbito de esta tesis, sería posible paralelizar el proceso de restauración y hacer uso de computadores masivamente paralelos o GPUs para procesar cada una de las parejas de bloques. El tamaño del bloque $B_x \times B_y$ en las direcciones horizontal y vertical a partir del número de bloques en horizontal y vertical, p_x y p_y , respectivamente, y del factor de solapamiento entre bloques adyacentes, $overlap$, $0 \leq overlap < 1$, se obtiene como

$$B_z = \left\lfloor \frac{N_z}{(1 - overlap)p_z + overlap} \right\rfloor, z \in \{x, y\},$$

con N_x y N_y el tamaño de las imágenes en horizontal y vertical, y $\lfloor \cdot \rfloor$ representa el operador de redondeo al entero más próximo por debajo. Por ejemplo, en una imagen de tamaño $N_y = 800$,

$N_x = 600$ píxeles, con número de parches $p_y = 8$ y $p_x = 6$ y un porcentaje de solapamiento del 50% ($overlap = 0,5$), obtendríamos un tamaño de bloque de $B_y = 177$ y $B_x = 171$ píxeles.

5.3. Estimación del núcleo de emborronamiento

Después de los procesos de calibración y división en bloques de ambas observaciones, podemos proceder a la estimación de la PSF a partir de cada una de las parejas de bloques. En este trabajo asumimos un modelo de degradación lineal y espacialmente invariante dentro del bloque, como en el capítulo previo, de modo que el proceso de observación de los parches de cada una de las imágenes puede ser expresado matemáticamente en forma de notación matriz-vector como

$$\mathbf{y}_1^p = \mathbf{H}^p \mathbf{x}^p + \mathbf{n}_1^p, \quad (5.1)$$

$$\mathbf{y}_2^p = \mathbf{x}^p + \mathbf{n}_2^p, \quad (5.2)$$

donde \mathbf{y}_1^p e \mathbf{y}_2^p son el parche p , $p \in (p1, p2)$ con $1 \leq p_1 \leq p_y$ y $1 \leq p_2 \leq p_x$, de las imágenes de larga y corta exposición, respectivamente, representado en forma de vector columna de tamaño $(B_y \times B_x) \times 1$. El parche de la imagen desconocida lo representaremos por \mathbf{x}^p , mientras que la matriz de convolución del emborronamiento espacialmente invariante, de dimensiones $(B_y \times B_x) \times (B_y \times B_x)$, la expresamos como \mathbf{H}^p . Finalmente \mathbf{n}_1^p y \mathbf{n}_2^p son las componentes de ruido blanco Gaussiano de media cero.

El emborronamiento se estima en cada parche para cada uno de los bloques observados \mathbf{y}_1^p e \mathbf{y}_2^p . Así, la matriz \mathbf{Y}_2^p representa el bloque de imagen \mathbf{y}_2^p escrita como matriz de convolución de dimensiones $(B_y \times B_x) \times (h_y \times h_x)$, con un tamaño del núcleo de emborronamiento \mathbf{h}^p , fijado por el usuario, de dimensiones (h_y, h_x) .

Usando (5.1) y (5.2), la estimación $\hat{\mathbf{h}}^p$, de un núcleo \mathbf{h}^p se obtiene como

$$\hat{\mathbf{h}}^p = \arg \min_{\mathbf{h}^p} \|\mathbf{y}_1^p - \mathbf{Y}_2^p \mathbf{h}^p\|^2, \quad (5.3)$$

sujeta a las restricciones $0 \leq \mathbf{h}_i^p \leq 1$ y $\sum_i \mathbf{h}_i^p = 1$, $i = 1, \dots, (h_x \times h_y)$. Este problema puede resolverse fácilmente como un problema lineal de mínimos cuadrados con restricciones. Con este fin se ha utilizado la rutina `lsqlin` de MATLAB. Claramente se produce un problema cuando deseamos calcular el valor de la convolución en las fronteras de los bloques. La solución consiste en ampliar el tamaño del bloque usando los valores de los píxeles de la imagen en bloques vecinos o replicar los valores de los bordes donde no haya parches vecinos disponibles.

La estimación del emborronamiento que hacemos en esta fase podría fallar en parches con escasa textura o con píxeles saturados, es decir, cuando no hay suficiente información disponible para la estimación del kernel. Las estimaciones pueden ser también poco fiables en bloques dónde se combinen varios tipos de emborronamientos, como ocurre en bloques con objetos en movimiento o con objetos a diferentes distancias. Existen una serie de métodos para reparar las PSFs y hacer frente a estos problemas (véase, por ejemplo, [40] o [7]). El método de deconvolución que se propone en la siguiente sección 5.4.2 utiliza un algoritmo de reparación de PSF para favorecer al proceso de deconvolución; sin embargo, en el procedimiento descrito en la sección 5.5, en lugar de corregir las PSFs estimadas, se utiliza un enfoque alternativo, en el que se opta por mantener la estimación del núcleo para después sustituir, con una mayor seguridad, las restauraciones que fueron erróneas, sustituyéndolas por el resultado de la aplicación de un método de eliminación de ruido sobre la imagen de corto tiempo de exposición. De esta forma podemos proponer un método de selección entre deconvolución o la eliminación de ruido más robusto (sección 5.5.1) que en el primero de los algoritmos (sección 5.4.2). Nótese que los dos métodos que proponemos en este capítulo hacen uso de la idea intuitiva subyacente por la que se elige entre la restauración proporcionada de la deconvolución mediante combinación de las dos observaciones (borrosa y ruidosa) o la eliminación de ruido exclusivamente sobre la imagen ruidosa. La aplicación de un método de eliminación de ruido a un único parche ruidoso puede proporcionar una mejor restauración que la obtenida por el proceso de deconvolución, ya que en bloques donde la estimación del emborronamiento fue poco precisa, por ejemplo por la presencia de un objeto en movimiento sobre fondo estático, la deconvolución podría detectar la imprecisión de la información proporcionada por la PSF y converger hacia una versión suavizada de la observación ruidosa, y en este caso el uso de un método de eliminación de ruido específico probablemente mejore la solución obtenida.

5.4. Deconvolución rápida

Una vez estimado el emborronamiento de cada bloque se procedería al proceso de deconvolución. En esta sección, proponemos un método de deconvolución novedoso¹, que combina pares de bloques borrosos de exposición correcta con los de corta exposición pero ruidosos para ob-

¹El algoritmo de restauración fue presentado en la 19th European Signal Processing Conference (EUSIPCO) de 2011 [7]

tener un conjunto de bloques nítidos sin emborronamiento ni ruido. La novedad del método propuesto reside en que una vez estimada la PSF de cada bloque, se realiza un proceso de reparación de PSFs, que se detalla en 5.4.1, con el fin de recuperar las PSFs en bloques donde no fue posible realizar una estimación precisa como, por ejemplo, bloques con poca textura o con emborronamientos distintos dentro del mismo bloque. Después de obtener el conjunto de PSFs reparadas se procede a estimar la imagen nítida desconocida empleando para ello un método automático que aplica selectivamente una deconvolución de ambos pares de bloques o elimina ruido del bloque de corta exposición y describiremos en la sección 5.4.2.

A continuación resumimos el Algoritmo 3 de Restauración Espacialmente Variante por Deconvolución Rápida (AREVDR). Cada una de las etapas de las que consta se describen en la sección correspondiente.

Algoritmo 3: Algoritmo de restauración espacialmente variante por deconvolución rápida (AREVDR)

Preprocesamiento: Calibración del par de imágenes y división en bloques (Sec. 5.2)

for cada par de bloques **do**

 | Estimación del núcleo de emborronamiento (Sec. 5.3).

end

Reparación del emborronamiento. Sección 5.4.1

for cada par de bloques **do**

 | Deconvolución/supresión de ruido del bloque. Sec. 5.4.2)

end

Posprocesamiento: Unir los bloques restaurados para conformar la imagen final. (Sec. 5.6)

5.4.1. Etapa de reparación de la PSF

En esta etapa, aplicamos una corrección a los núcleos estimados en bloques con escasa textura o con píxeles saturados, es decir cuando no hay suficiente información disponible para la estimación del emborronamiento. El algoritmo 4 corrige la PSF estimada en el parche si la diferencia con los núcleos en bloques vecinos es superior a un umbral, y luego la sustituye por el promedio de los núcleos de bloques adyacentes incluyendo la estimación del propio bloque. Hemos comprobado que el umbral puede ajustarse de forma adecuada a la mitad del rango de las diferencias con sus vecinos. Un procedimiento alternativo para la corrección del núcleo se propone en [40]. El objetivo de esta etapa va a ser el de favorecer una deconvolución correcta

de los pares de bloques evitando abusar del método de eliminación de ruido que debe ser usado sólo en bloques que son irrecuperables mediante una deconvolución.

Algoritmo 4: Algoritmo de corrección de PSFs.

Input: El emborronamiento espacialmente variante expresado en forma matricial con un núcleo \mathbf{h}^p en cada parche p

Output: La matriz de PSFs espacialmente variante corregida.

for Para cada núcleo \mathbf{h}^p **do**

Calcular la PSF promedio \mathbf{h}_{mean}^p como la media de los núcleos de los bloques vecinos de \mathbf{h}^p , incluyendo \mathbf{h}^p .

Calcular $diff^p = \|\mathbf{h}^p - \mathbf{h}_{mean}^p\|_1$.

end

Establecer el umbral $thr = \frac{\max(diff) + \min(diff)}{2} + \min(diff)$, donde $diff$ representa el vector formado por $diff^p, p = 1, \dots, (p_x \times p_y)$

Para cada PSF \mathbf{h}^p tal que $diff^p > thr$, sustituir \mathbf{h}^p por \mathbf{h}_{mean}^p .

5.4.2. Deconvolución

Después de la estimación del núcleo y la etapa de corrección, estamos en disposición de aplicar el método de deconvolución para obtener una estimación precisa de la imagen nítida en cada bloque. A continuación, describimos el proceso de deconvolución que aplicamos sobre cada uno de los pares de bloques. Para ello hacemos uso del algoritmo de deconvolución Bayesiano desarrollado en [2] y aplicamos una serie de modificaciones orientadas al ahorro de tiempo y memoria. En este punto nos permitimos omitir los detalles del algoritmo de deconvolución, proporcionando exclusivamente las modificaciones a que fue sometido y los resultados obtenidos.

La idea clave en [2] consiste en modelar las incógnitas dentro de un marco Bayesiano jerárquico y desarrollar un algoritmo de deconvolución ciega que estima, de forma conjunta, la imagen desconocida y el emborronamiento. Sin embargo, la estimación que obtenemos directamente de aplicar [2] no proporciona resultados aceptables en nuestro caso, debido al reducido tamaño de los parches que hace inviable la estimación conjunta de la imagen y la PSF. Sin embargo, puesto que el núcleo que estimamos siguiendo la sección 5.3 y el algoritmo 4 es bastante preciso, proponemos utilizar el método desarrollado en [2] pero utilizando un modelo más simple en el que previamente estimamos el núcleo de emborronamiento y posteriormente lo utilizamos directamente en el proceso de estimación de la imagen nítida.

Para este algoritmo proponemos emplear un enfoque variacional equivalente al empleado en [2] para obtener una aproximación de la distribución a posteriori de la imagen restaurada y los parámetros pero utilizando un proceso iterativo alternativo a la técnica del gradiente conjugado, de la siguiente forma: partiendo de las ecuaciones (5.1) y (5.2), asumimos que los ruidos \mathbf{n}_1^p y \mathbf{n}_2^p en las dos observaciones siguen distribuciones Gaussianas independientes de media cero y varianzas $(\beta_1^p)^{-1}$, $(\beta_2^p)^{-1}$, respectivamente. Proponemos el uso, además, de un modelo a priori TV sobre la imagen y distribuciones uniformes, no informativas, para modelar el conocimiento a priori sobre los parámetros del modelo de imagen y de ruido. La estimación de la imagen para cada parche p la obtenemos mediante

$$\mathbf{x}^p = \Sigma_x^p (\beta_1^p (\mathbf{H}^p)^t \mathbf{y}_1^p + \beta_2^p \mathbf{y}_2^p) \quad (5.4)$$

$$(\Sigma_x^p)^{-1} = \beta_1^p (\mathbf{H}^p)^t \mathbf{H}^p + \alpha^p ((\Delta^u)^t \mathbf{W}^p \Delta^u + (\Delta^v)^t \mathbf{W}^p \Delta^v) + \beta_2^p \mathbf{I}, \quad (5.5)$$

donde Δ^u y Δ^v representan las aproximaciones al operador gradiente en las direcciones horizontal y vertical, respectivamente, el término \mathbf{W}^p representa la matriz de adaptividad espacial que controla el nivel de suavizado en cada píxel dependiendo de la variación de intensidad en la posición del píxel, según indique el vector \mathbf{w} el cual representa la variación total de la imagen estimada y controla el equilibrio entre las fronteras de los objetos y el suavizado de la imagen estimada. β_1^p y β_2^p denotan las inversas de las varianzas del ruido correspondiente a las observaciones \mathbf{y}_1^p e \mathbf{y}_2^p , respectivamente. Nótese que las ecuaciones 5.4 y 5.5 coinciden con las descritas en el capítulo 4 para la estimación de la esperanza de la imagen \mathbf{x} , ecuaciones 4.42 y 4.43, con $\lambda_1 = \lambda_2 = 1$.

La estimación de cada uno de los parámetros se realiza siguiendo las siguientes ecuaciones

$$w_j^p = (\Delta_j^u(\mathbf{x}^p))^2 + (\Delta_j^v(\mathbf{x}^p))^2, j = 1, \dots, B \quad (5.6)$$

$$\mathbf{W}^p = \text{diag} \left[\frac{1}{\sqrt{w_j^p} + \epsilon} \right], j = 1, \dots, B \quad (5.7)$$

$$\alpha^p = \frac{B}{2 \sum_j \sqrt{w_j^p}}, j = 1, \dots, B \quad (5.8)$$

$$\beta_1^p = \frac{B}{\|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}^p\|^2} \quad (5.9)$$

$$\beta_2^p = \frac{B}{\|\mathbf{y}_2^p - \mathbf{x}^p\|^2}, \quad (5.10)$$

con $B = B_x \times B_y$, el número de pixels en el bloque. Nótese que en este método simplificamos el cálculo de la estimación de parámetros en [2], ya que asumimos que la distribución \mathbf{x}^p es degenerada y, por tanto, las esperanzas que aparecen originalmente en el artículo [2] y que coinciden con las Ecs. 5.6, 5.7 y 5.8 a 5.10 para el caso de una distribución no informativa sobre los parámetros, pueden ser calculadas sin tener en cuenta la contribución de las trazas correspondientes.

Para calcular \mathbf{x}^p en la ecuación (5.4) y los parámetros de las ecuaciones (5.6)-(5.10) utilizamos el siguiente proceso iterativo: Inicialmente establecemos como estimación inicial de \mathbf{x}^p , el bloque de larga exposición \mathbf{y}_1^p . La actualización de los parámetros se realiza según las ecuaciones (5.6)-(5.10), usamos ϵ , un valor positivo muy pequeño, para evitar la división por cero en la ecuación 5.7. Una vez estimados los parámetros procedemos a calcular una nueva estimación de \mathbf{x}^p (ecuación 5.4), siguiendo el procedimiento que detallamos a continuación y volvemos a estimar los parámetros con el valor actualizado de \mathbf{x}^p .

En [2] se recurre al método de gradiente conjugado para evitar el cálculo explícito de la matriz de covarianza (Σ_x^p) de la Ec. (5.5). Debido al coste computacional que ello supondría en nuestro enfoque espacialmente variante, empleamos un procedimiento alternativo para la estimación de la imagen en cada bloque. Concretamente, resulta más eficiente emplear un método similar al de [80] con el fin de resolver la Ec. (5.4) de forma directa. En primer lugar reescribimos la Ec. (5.4) como

$$(\Sigma_x^p)^{-1} \mathbf{x}^p = \beta_1^p (\mathbf{H}^p)^t \mathbf{y}_1^p + \beta_2^p \mathbf{y}_2^p.$$

Nótese que en un determinado píxel j , se tiene

$$1 = \frac{(\beta_1^p (\mathbf{H}^p)^t \mathbf{y}_1^p + \beta_2^p \mathbf{y}_2^p)_j}{((\Sigma_x^p)^{-1} \mathbf{x}^p)_j}, j = 1, \dots, B. \quad (5.11)$$

Multiplicando ambas partes de la Ec. (5.11) por \mathbf{x}^p , nos permite definir un proceso iterativo para la actualización de la imagen.

$$(\mathbf{x}_{new}^p)_j = (\mathbf{x}_{old}^p)_j \times \frac{(\beta_1^p (\mathbf{H}^p)^t \mathbf{y}_1^p + \beta_2^p \mathbf{y}_2^p)_j}{(\beta_1^p (\mathbf{H}^p)^t \mathbf{H}^p \mathbf{x}_{old}^p + \alpha^p [\mathbf{D}^t \mathbf{Z}^p \mathbf{D}] \mathbf{x}_{old}^p + \beta_2^p \mathbf{x}_{old}^p)_j}, \quad (5.12)$$

donde $\mathbf{D}^t = [(\Delta^u)^t, (\Delta^v)^t]$, $\mathbf{Z}^p = \begin{bmatrix} \mathbf{W}^p & 0 \\ 0 & \mathbf{W}^p \end{bmatrix}$, \mathbf{x}_{new}^p denota la nueva estimación de \mathbf{x}^p y \mathbf{x}_{old}^p representa la estimación de \mathbf{x}^p antes de la actualización de parámetros. Todos los parámetros en la parte derecha de la Ec. (5.12) se calculan utilizando \mathbf{x}_{old}^p en las Ecs. (5.6-5.10).

En la sección 5.7 se han llevado a cabo una serie de experimentos sintéticos y reales con distintos algoritmos de restauración tales como Richardson-Lucy [29, 30], y el estado del arte en exposición dual [2, 3, 4], pero en general producen imágenes menos nítidas que el propuesto, en especial las restauraciones de RL con un considerable efecto de anillado.

Algoritmo 5: Algoritmo de deconvolución propuesto

Input: $\mathbf{y}_1^p, \mathbf{y}_2^p$ los bloques de larga y corta exposición respectivamente; \mathbf{h}^p la PSF corregida para el parche p .

Output: \mathbf{x}^p el parche restaurado.

$\mathbf{x}_{new}^p = \mathbf{y}_1^p$.

repeat

$\mathbf{x}_{old}^p = \mathbf{x}_{new}^p$.

 Encontrar la estimación de los parámetros usando las Ecs. (5.6)-(5.10) con $\mathbf{x}^p = \mathbf{x}_{old}^p$.

 Encontrar \mathbf{x}_{new}^p , la nueva estimación del bloque de imagen usando la Ec. (5.12).

until $\frac{\|\mathbf{x}_{new}^p - \mathbf{x}_{old}^p\|^2}{\|\mathbf{x}_{old}^p\|^2} > threshold$

$\mathbf{x}_1^p = \mathbf{x}_{new}^p$, el bloque resultado de la deconvolución.

Obtener \mathbf{x}_2^p de aplicar algoritmo de eliminación de ruido en [1] usando como estimación de la varianza $(\beta_2^p)^{-1}$.

Calcular medidas de error \mathbf{x}_1^p y \mathbf{x}_2^p usando la Ec. (5.13).

if ($error_{12} > errThreshold$ AND $error_{21} > errThreshold$) OR $error_{11} > error_{12}$ **then**

$\mathbf{x}^p = \mathbf{x}_2^p$.

else

$\mathbf{x}^p = \mathbf{x}_1^p$.

end

Dotar de robustez a PSFs incorrectas

El método de deconvolución propuesto proporciona imágenes de alta calidad para emborronamiento por movimiento de cámara; sin embargo, al igual que otros algoritmos de deconvolución, la estimación de imagen obtenida con nuestro método no es suficiente cuando la escena contiene objetos en movimiento. Observe que el modelo $\mathbf{y}_1^p = \mathbf{H}^p \mathbf{x}^p + \mathbf{n}_1^p$, es poco realista en caso de emborronamiento por objeto en movimiento, ya que este tipo de emborronamiento es muy probable que sea variante en el parche, además el algoritmo tiene que diferenciar los píxeles de fondo (donde el emborronamiento es invariante) del objeto u objetos en primer plano con

otro emborronamiento distinto. Esto podría lograrse mediante el uso de un algoritmo de segmentación, que en nuestro caso resulta inviable debido al coste computacional que supondría. Recordemos que en el marco propuesto, trabajamos sobre una colección de bloques en la dividimos la imagen, la idea consiste en seleccionar en cada uno de estos parches si utilizamos la restauración que nos proporciona el algoritmo de deconvolución o una versión con supresión de ruido del parche de baja exposición. Observe que, aprovechando el corto tiempo de exposición de \mathbf{y}_2 , esperamos que esta observación no se vea afectada por emborronamiento por movimiento. Por lo tanto, en parches afectados por este tipo de emborronamiento, esperamos obtener mejores resultados con la eliminación de ruido en \mathbf{y}_2^p que combinando ambas observaciones. De este modo, si la estimación del núcleo no fuese correcta, podemos tratar de eliminar el ruido exclusivamente en el parche de la segunda observación \mathbf{y}_2^p y sustituir el resultado de la deconvolución por este otro.

Inicialmente obtenemos dos restauraciones en cada bloque, la primera de ellas es el resultado de aplicar nuestro algoritmo de deconvolución combinando información de las dos observaciones. La segunda se obtiene directamente de aplicar el método de eliminación de ruido en [1] sobre la imagen de corta exposición. El método en [1] necesita una estimación de la varianza del ruido. Dado que nuestro algoritmo de deconvolución ya estimó los parámetros del ruido, vamos a utilizar la inversa del valor calculado usando la ecuación (5.10). Las dos restauraciones obtenidas por el algoritmo de deconvolución y el de eliminación de ruido las denotaremos por \mathbf{x}_1^p y \mathbf{x}_2^p , respectivamente. Para seleccionar una de las restauraciones calculamos las siguientes medidas de error

$$\begin{aligned} error_{1k} &= \|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}_k^p\|^2, \\ error_{2k} &= \|\mathbf{y}_2^p - \mathbf{x}_k^p\|^2, \end{aligned} \quad (5.13)$$

para $k = 1, 2$.

Finalmente el bloque resultante del método de eliminación de ruido se selecciona en caso que ambos errores ($error_{12}$ y $error_{21}$) sean mayores que cierto umbral ($errThreshold = error_{11} + error_{22}$), en otro caso, seleccionamos el parche resultado de la deconvolución. La justificación de esta estrategia de selección se origina del hecho que en caso de objeto en movimiento, los núcleos estimados contienen errores debido a la mezcla de partes estáticas con otras en movimiento, y por tanto, una única PSF en el bloque no puede explicar ambas situaciones. Por

ello, el error $error_{12}$ tendrá valores grandes. Además, debido a los errores de deconvolución, la imagen deconvolucionada presenta una baja semejanza con la imagen ruidosa, lo que implica que $error_{21}$ sea elevado. En este caso, la información aportada exclusivamente por la imagen ruidosa, tras eliminación de ruido, es suficiente para obtener una restauración correcta del bloque. Además, la imagen resultado de eliminación de ruido también se selecciona en el caso que el error de ésta sea menor que la restauración proporcionada por el método de deconvolución, es decir ($error_{11} > error_{12}$), situación que rara vez ocurre en la práctica.

El tiempo de ejecución del método de deconvolución propuesto y resumido en el Algoritmo 5 es muy ajustado, siendo de menos de 2 minutos en un ordenador portátil con procesador de doble núcleo de 2,13 Ghz y 4GB de memoria RAM; para un tamaño de imagen de 512×512 , un tamaño de 47×47 píxeles para la función de emborronamiento, división en 6×6 bloques. Nótese que la implementación podría ser mejorada notablemente explotando el paralelismo y ejecución en GPU.

5.5. Deconvolución basada en combinación de modelos de imagen

Si bien el algoritmo descrito en la Sección 5.4 proporciona una deconvolución rápida, el uso de la combinación de modelos a priori de una manera similar a la descrita en el capítulo anterior permite obtener una deconvolución más precisa. Por ello en esta sección presentamos un procedimiento ² Bayesiano automático que combina dos modelos a priori sobre la imagen para estimar la imagen desconocida haciendo uso de los núcleos estimados previamente y el par de imágenes observadas.

Al igual que el método que hemos descrito previamente en la Sección 5.4, cuando la estimación de la PSF no es suficientemente precisa debido a, por ejemplo, movimiento de objetos o ausencia de texturas en el parche, el resultado de la deconvolución del bloque podría no ser correcto. Para detectar y corregir estos casos, proponemos una estrategia de selección, alternativa a la descrita en la Sección 5.4.2, en la que seleccionamos la versión de eliminación de ruido del parche de corta exposición en bloques donde la deconvolución no resultó ser suficiente. Los

²Los resultados preliminares de este algoritmo de restauración fueron presentados en el 7th International Symposium on Image and Signal Processing and Analysis de 2011[6] y han sido publicados recientemente en la revista Information Fusion[9].

resultados experimentales, presentados en la Sección 5.7, demuestran que el método propuesto ofrece una estimación del núcleo espacialmente variante correcta y una imagen de alta calidad incluso en conjuntos de imágenes altamente degradadas.

A continuación resumimos el algoritmo de restauración propuesto para el método de deconvolución que tratamos en esta sección, cada una de las etapas se detalla en su respectiva sección. Nótese que, a diferencia del Algoritmo 3 (AREVDR) propuesto en la Sección 5.4, en este algoritmo, la estimación del núcleo y la imagen restaurada se realiza de forma independiente para cada uno de los bloques. Esto da lugar a un algoritmo que puede ser computado de forma eficiente en procesadores masivamente paralelos o GPUs.

Algoritmo 6: Algoritmo de restauración espacialmente variante. con deconvolución basada en combinación de modelos (AREVDCM)

Preprocesamiento: Calibración del par de imágenes y división en bloques (Sec. 5.2)

for *cada bloque* **do**

 Estimación del núcleo (Sec. 5.3)

 Deconvolución/supresión de ruido del bloque. Sec. 5.5.1)

end

Posprocesamiento: Fusión de los bloques restaurados para conformar la imagen final. (Sec. 5.6)

5.5.1. Deconvolución

Una vez estimado el núcleo para el bloque considerado siguiendo el procedimiento descrito en la Sección 5.3, buscamos la mejor estimación de la imagen nítida de dicho bloque. Para ello se ha desarrollado un algoritmo de deconvolución usando una formulación Bayesiana jerárquica para el problema de exposición dual. La idea fundamental reside en la combinación de un modelo raro (sparse) de imagen (en términos de abundancia de ceros) con otro no raro. En concreto, combinamos el modelo de variación total (TV) [74] con el simultáneo autorregresivo (SAR) [81]. La idea de la combinación de modelos ha sido utilizado con éxito también en otras áreas, como la restauración de imágenes astronómicas [20], super-resolución [82] o la descomposición de la imagen [83].

El objetivo de la combinación de modelos de imagen es recuperar tantos detalles como sea posible de la imagen nítida desconocida a partir de las dos observaciones y las bondades de cada uno de los modelos de imagen TV y SAR. El modelo de imagen TV tiene la capacidad de

preservar las fronteras de los objetos, pero tiende a suavizar las áreas con textura, mientras que el modelo SAR conserva las transiciones de texturas mejor, pero suaviza en exceso las fronteras. Nuestro objetivo es preservar tanto la textura y las fronteras de la imagen latente mediante la incorporación de ambos modelos.

Modelo Bayesiano jerárquico

A continuación describimos las distribuciones a priori utilizadas para modelar cada una de las variables desconocidas.

La distribución a priori sobre la imagen del modelo TV (ralo) viene dada por

$$p_1(\mathbf{x}^p | \alpha_1^p) \propto (\alpha_1^p)^{B/2} \exp \left[-\alpha_1^p \sum_{j=1}^B \sqrt{(\Delta_j^u(\mathbf{x}^p))^2 + (\Delta_j^v(\mathbf{x}^p))^2} \right], \quad (5.14)$$

donde $j = 1, \dots, B$ con $B = B_y \times B_x$ representa el número de pixels en el bloque, α_1^p es el parámetro del modelo, y los operadores $\Delta_j^u(\mathbf{x})$ y $\Delta_j^v(\mathbf{x})$ corresponden, respectivamente, a las diferencias de primer orden en horizontal y vertical para el pixel j .

El segundo modelo de imagen que combinamos con el anterior, modelo SAR (no ralo), se define como

$$p_2(\mathbf{x}^p | \alpha_2^p) \propto (\alpha_2^p)^{B/2} \exp \left[-\frac{\alpha_2^p}{2} \|\mathbf{C}\mathbf{x}^p\|^2 \right], \quad (5.15)$$

donde \mathbf{C} representa el operador Laplaciano y α_2^p es el parámetro del modelo. Nótese que en un principio podríamos haber supuesto un único modelo a priori sobre la imagen, combinación de los dos anteriores, de la siguiente forma

$$p(\mathbf{x}^p | \alpha_1^p, \alpha_2^p) = \frac{1}{Z(\alpha_1^p, \alpha_2^p)} \exp \left[-\alpha_1^p \sum_{j=1}^B \sqrt{(\Delta_j^u(\mathbf{x}^p))^2 + (\Delta_j^v(\mathbf{x}^p))^2} - \frac{\alpha_2^p}{2} \|\mathbf{C}\mathbf{x}^p\|^2 \right]. \quad (5.16)$$

Sin embargo, la estimación de los parámetros para este modelo es muy compleja, debido a que la función de partición $Z(\alpha_1^p, \alpha_2^p)$ es desconocida y difícil de aproximar.

A partir de los modelos (5.1) and (5.2), asumiendo que los ruidos \mathbf{n}_1^p y \mathbf{n}_2^p siguen una distribución Gaussiana de media cero y varianzas $(\beta_1^p)^{-1}$ y $(\beta_2^p)^{-1}$, respectivamente, podemos escribir la distribución de probabilidad condicionada de las observaciones como

$$p(\mathbf{y}_1^p, \mathbf{y}_2^p | \mathbf{x}^p, \beta_1^p, \beta_2^p) \propto (\beta_1^p)^{B/2} (\beta_2^p)^{B/2} \exp \left[-\frac{\beta_1^p}{2} \|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}^p\|^2 - \frac{\beta_2^p}{2} \|\mathbf{y}_2^p - \mathbf{x}^p\|^2 \right]. \quad (5.17)$$

La combinación de distribuciones a priori sobre la imagen y modelos de observación dependen de un conjunto de parámetros cuyos valores han de ser estimados conjuntamente con

la imagen. Para su modelado empleamos distribuciones a priori impropias, no informativas, $p(\xi) \propto \text{const}$, $\xi \in \{\alpha_1^p, \alpha_2^p, \beta_1^p, \beta_2^p\}$, definidas en $(0, \infty)$.

Finalmente, obtenemos una distribución conjunta para cada uno de los modelos a priori de imagen en las Ecs. (5.14) y (5.15), y la combinación del modelo de observación (Ec. 5.17) y las a priori no informativas para los parámetros de los modelos de observación e imagen como

$$p_1(\Omega, \alpha_1, \mathbf{y}_1^p, \mathbf{y}_2^p) = p(\mathbf{x}^p | \alpha_1^p) p(\alpha_1^p) p(\mathbf{y}_1^p, \mathbf{y}_2^p | \Omega) p(\beta_1^p) p(\beta_2^p), \quad (5.18)$$

$$p_2(\Omega, \alpha_2, \mathbf{y}_1^p, \mathbf{y}_2^p) = p(\mathbf{x}^p | \alpha_2^p) p(\alpha_2^p) p(\mathbf{y}_1^p, \mathbf{y}_2^p | \Omega) p(\beta_1^p) p(\beta_2^p), \quad (5.19)$$

donde $\Omega = \{\mathbf{x}^p, \beta_1^p, \beta_2^p\}$. Los parches de imagen desconocidos \mathbf{x}^p , al igual que los parámetros del modelo, los vamos a estimar a partir de la distribución conjunta.

Inferencia Bayesiana Variacional

En esta sección, empleamos un enfoque variacional para obtener una aproximación de la distribución a posteriori del bloque de la imagen y los parámetros, que describimos a continuación. Denotemos por Θ el conjunto de incógnitas, es decir, $\Theta = \{\Omega, \alpha_1^p, \alpha_2^p\}$. El objetivo es aproximar la distribución a posteriori de $p(\Theta | \mathbf{y}_1^p, \mathbf{y}_2^p)$ por otra distribución $q(\Theta)$ que permita un análisis más sencillo. En general, la única suposición en el análisis Bayesiano variacional es que la distribución $q(\Theta)$ es factorizable [54]. Concretamente hacemos uso de la siguiente factorización:

$$q(\Theta) = q(\Omega) q(\alpha_1^p) q(\alpha_2^p), \quad (5.20)$$

with $q(\Omega) = q(\mathbf{x}^p) q(\beta_1^p) q(\beta_2^p)$. La distribución aproximada $q(\Theta)$ la encontramos como la minimización de la divergencia de Kullback-Leibler (KL) entre $q(\Theta)$ y $p(\Theta | \mathbf{y}_1^p, \mathbf{y}_2^p)$, expresada como

$$\hat{q}(\Theta) = \underset{q(\Theta)}{\operatorname{argmin}} \sum_{k=1}^2 \lambda_k C_{KL}(q(\Omega) q(\alpha_k^p) \| p_k(\Omega, \alpha_k^p | \mathbf{y}_1^p, \mathbf{y}_2^p)), \quad (5.21)$$

con $\lambda_k \geq 0$, $\lambda_1 + \lambda_2 = 1$, y la divergencia KL [57], para $k = 1, 2$, como

$$C_{KL}(q(\Omega) q(\alpha_k^p) \| p_k(\Omega, \alpha_k^p | \mathbf{y}_1^p, \mathbf{y}_2^p)) = \int q(\Theta) \log \left(\frac{q(\Omega) q(\alpha_k^p)}{p_k(\Omega, \alpha_k^p | \mathbf{y}_1^p, \mathbf{y}_2^p)} \right) d\Theta + \text{const}, \text{ para } k = 1, 2. \quad (5.22)$$

La estimación de los parámetros λ_1 y λ_2 no se llevará a cabo en esta memoria pero demostraremos empíricamente que la combinación no degenerada de divergencias ($0 < \lambda_1 < 1$) produce mejores resultados que estableciendo $\lambda_1 = 0$ o $\lambda_1 = 1$.

La expresión (5.21) puede reescribirse como

$$\hat{q}(\Theta) = \operatorname{argmin}_{q(\Theta)} \int q(\Theta) \times \log \left(\frac{q(\Omega)}{p(\mathbf{y}_1^p, \mathbf{y}_2^p | \Omega) p(\beta_1^p) p(\beta_2^p)} \prod_{k=1}^2 \left[\frac{q(\alpha_k^p)}{p_k(\mathbf{x}^p | \alpha_k^p) p(\alpha_k^p)} \right]^{\lambda_k} \right) d\Theta. \quad (5.23)$$

Por desgracia, el análisis Bayesiano variacional no puede ser utilizado directamente; ya que la raíz cuadrada del modelo TV utilizado como a priori sobre la imagen complica la solución de la integral en (5.23). Los problemas causados por el modelo a priori TV pueden ser evitados mediante la utilización de un enfoque de mayorización-minimización (MM), cuyos detalles se muestran en [74]. Este método encuentra una cota para la distribución en la Ec. (5.14) y hace posible la derivación analítica de la inferencia Bayesiana, que mostramos en las líneas siguientes.

Consideremos el funcional $\mathbf{M}(\alpha_1^p, \mathbf{x}^p, \mathbf{w}^p)$, donde $\mathbf{w}^p \in (R^+)^B$ es un vector B -dimensional de componentes w_j , $j = 1, \dots, B$,

$$\mathbf{M}(\alpha_1^p, \mathbf{x}^p, \mathbf{w}^p) = (\alpha_1^p)^{B/2} \exp \left[-\frac{\alpha_1^p}{2} \sum_{j=1}^B \frac{(\Delta_j^u(\mathbf{x}^p))^2 + (\Delta_j^v(\mathbf{x}^p))^2 + w_j}{\sqrt{w_j}} \right]. \quad (5.24)$$

El vector auxiliar \mathbf{w}^p necesita calcularse simultáneamente con la imagen \mathbf{y} , como ya comentamos en la Sección 5.4.2, tiene una relación directa con los gradientes del bloque de la imagen desconocida \mathbf{x}^p . Como ya demostramos en la Sección 4.4, el funcional $\mathbf{M}(\alpha_1^p, \mathbf{x}^p, \mathbf{w}^p)$ es una cota inferior de la distribución a priori de la imagen $p_1(\mathbf{x}^p | \alpha_1^p)$, esto es,

$$p_1(\mathbf{x}^p | \alpha_1^p) \geq \mathbf{M}(\alpha_1^p, \mathbf{x}^p, \mathbf{w}^p). \quad (5.25)$$

Usando la Ec. (5.25), la cota inferior de la distribución conjunta de probabilidad en la Ec. (5.18), se demuestra que

$$\begin{aligned} p_1(\Omega, \alpha_1^p, \mathbf{y}_1^p, \mathbf{y}_2^p) &\geq \mathbf{M}(\alpha_1^p, \mathbf{x}^p, \mathbf{w}^p) p(\alpha_1^p) p(\mathbf{y}_1^p, \mathbf{y}_2^p | \Omega) p(\beta_1^p) p(\beta_2^p) \\ &= \mathbf{F}(\Omega, \alpha_1^p, \mathbf{w}^p, \mathbf{y}_1^p, \mathbf{y}_2^p), \end{aligned} \quad (5.26)$$

que lleva a la siguiente cota superior de la divergencia de KL en la Ec. (5.22)

$$C_{KL}(q(\Omega)q(\alpha_1^p) \parallel p_1(\Omega, \alpha_1^p, \mathbf{y}_1^p, \mathbf{y}_2^p)) \leq \min_{\mathbf{w}^p} C_{KL}(q(\Omega)q(\alpha_1^p) \parallel \mathbf{F}(\Omega, \alpha_1^p, \mathbf{w}^p, \mathbf{y}_1^p, \mathbf{y}_2^p)). \quad (5.27)$$

Antes de proceder al cálculo de la aproximación de la distribución a posteriori, observamos que para calcular $q(\alpha_k)$, con $k = 1, 2$, únicamente tenemos que tener en cuenta la divergencia

donde aparece esa distribución. Por tanto, podemos escribir

$$q(\alpha_1^p) = \text{const} \times \exp(\langle \log \mathbf{F}(\Omega, \alpha_1^p, \mathbf{w}^p, \mathbf{y}_1^p, \mathbf{y}_2^p) \rangle_\Omega), \quad (5.28)$$

$$q(\alpha_2^p) = \text{const} \times \exp(\langle \log p_2(\Omega, \alpha_2^p, \mathbf{y}_1^p, \mathbf{y}_2^p) \rangle_\Omega), \quad (5.29)$$

donde $\langle \cdot \rangle_\Omega = \mathbf{E}_{q(\Omega)}[\cdot]$, la esperanza con respecto a la distribución $q(\Omega)$.

Además, para calcular el resto de distribuciones, $q(\omega)$, $\omega \in \Omega$, tenemos que tener en cuenta ambas divergencias, obteniendo

$$q(\omega) = \text{const} \times \exp \left(\left\langle \log p(\mathbf{y}_1^p, \mathbf{y}_2^p | \Omega) p(\beta_1^p) p(\beta_2^p) \times \left[\mathbf{M}(\alpha_1^p, \mathbf{x}^p, \mathbf{w}^p) p(\alpha_1^p) \right]^\lambda \left[p_2(\mathbf{x}^p | \alpha_2^p) p(\alpha_2^p) \right]^{1-\lambda} \right\rangle_{\Omega_\omega} \right), \quad (5.30)$$

donde Ω_ω denota el conjunto de incógnitas con ω eliminado, y $\lambda = \lambda_1$, por tanto, $\lambda_2 = 1 - \lambda$.

Cálculo de las aproximaciones de la distribución a posteriori

El cálculo de las distribuciones anteriores para cada una de las incógnitas da como resultado un proceso iterativo, que converge a la mejor aproximación de la distribución a posteriori $p(\Theta | \mathbf{y}_1^p, \mathbf{y}_2^p)$ por distribuciones de la forma en la Ec. (5.21). En la presente sección, utilizamos las esperanzas de estas distribuciones, como las estimaciones puntuales de las incógnitas. La forma explícita de cada una de estas distribuciones se muestra a continuación.

La distribución $q(\mathbf{x}^p)$ se calcula a partir de la Ec. (5.30) como la distribución Gaussiana multivariante $q(\mathbf{x}^p) = \mathcal{N}(\mathbf{x}^p | \langle \mathbf{x}^p \rangle, \Sigma_{\mathbf{x}}^p)$ donde su media y matriz de covarianzas vienen dadas por

$$\langle \mathbf{x}^p \rangle = \Sigma_{\mathbf{x}}^p (\langle \beta_1^p \rangle (\mathbf{H}^p)^t \mathbf{y}_1^p + \langle \beta_2^p \rangle \mathbf{y}_2^p), \quad (5.31)$$

$$(\Sigma_{\mathbf{x}}^p)^{-1} = \lambda \langle \alpha_1^p \rangle ((\Delta^u)^t \mathbf{W}^p \Delta^u + (\Delta^v)^t \mathbf{W}^p \Delta^v) + (1 - \lambda) \langle \alpha_2^p \rangle \mathbf{C}^t \mathbf{C} + \langle \beta_1^p \rangle (\mathbf{H}^p)^t \mathbf{H}^p + \langle \beta_2^p \rangle \mathbf{I}, \quad (5.32)$$

donde Δ^u y Δ^v representan las aproximaciones discretas de los gradientes horizontales y verticales, respectivamente, \mathbf{W}^p es la matriz diagonal de tamaño $B \times B$ con la forma $\mathbf{W}^p = \text{diag} \left((w_j^p)^{-\frac{1}{2}} \right)$, para $j = 1, \dots, B$ con

$$w_j^p = \left\langle \left(\Delta_j^u(\mathbf{x}^p) \right)^2 + \left(\Delta_j^v(\mathbf{x}^p) \right)^2 \right\rangle, j = 1, \dots, B. \quad (5.33)$$

La matriz \mathbf{W}^p puede interpretarse como una matriz de ponderación espacialmente variante que proporciona adaptabilidad espacial al modelo, controlando la cantidad de suavizado en cada posición de pixel en función de la variación de la intensidad alrededor de este píxel. El parámetro λ controla la contribución de cada modelo de imagen. Por ejemplo, para un valor $\lambda =$

0 produce el modelo de restauración clásica SAR, mientras que $\lambda = 1$ conduce a la restauración del modelo TV. Una combinación razonable de ambos modelos permite recuperar las texturas de la imagen, mientras preservamos las fronteras [20].

El resto de parámetros se estiman como la esperanza de las distribuciones de las Ecs. (5.28), (5.29) y (5.30) dadas por

$$\langle \alpha_1^p \rangle = \frac{B}{2 \sum_{j=1}^B \sqrt{w_j^p}}, \quad (5.34)$$

$$\langle \alpha_2^p \rangle = \frac{B}{\langle \|\mathbf{C}\mathbf{x}^p\|^2 \rangle}, \quad (5.35)$$

$$\langle \beta_1^p \rangle = \frac{B}{\langle \|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}^p\|^2 \rangle}, \quad (5.36)$$

$$\langle \beta_2^p \rangle = \frac{B}{\langle \|\mathbf{y}_2^p - \mathbf{x}^p\|^2 \rangle}. \quad (5.37)$$

Las formas explícitas de las esperanzas en las Ecuaciones 5.33-5.37 se muestran a continuación

$$\left\langle \left(\Delta_j^u(\mathbf{x}^p) \right)^2 + \left(\Delta_j^v(\mathbf{x}^p) \right)^2 \right\rangle = \left(\Delta_j^u(\langle \mathbf{x}^p \rangle) \right)^2 + \left(\Delta_j^v(\langle \mathbf{x}^p \rangle) \right)^2 + \text{trace} \left(((\Delta^u)^t \Delta^u + (\Delta^v)^t \Delta^v) \Sigma_{\mathbf{x}}^p \right), \quad (5.38)$$

$$\langle \|\mathbf{C}\mathbf{x}^p\|^2 \rangle = \|\mathbf{C}\langle \mathbf{x}^p \rangle\|^2 + \text{trace}(\mathbf{C}^t \mathbf{C} \Sigma_{\mathbf{x}}^p), \quad (5.39)$$

$$\langle \|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}^p\|^2 \rangle = \|\mathbf{y}_1^p - \mathbf{H}^p \langle \mathbf{x}^p \rangle\|^2 + \text{trace}(\mathbf{H}^t \mathbf{H} \Sigma_{\mathbf{x}}^p), \quad (5.40)$$

$$\langle \|\mathbf{y}_2^p - \mathbf{x}^p\|^2 \rangle = \|\mathbf{y}_2^p - \langle \mathbf{x}^p \rangle\|^2 + \text{trace}(\Sigma_{\mathbf{x}}^p). \quad (5.41)$$

Para simplificar su cálculo asumimos, para la estimación de los parámetros, que la distribución \mathbf{x}^p es degenerada y, por tanto, las esperanzas en Ecs. (5.38)-(5.41) pueden ser calculadas sin tener en cuenta la contribución de las trazas correspondientes.

Para seleccionar el mejor valor de λ buscamos la estimación del bloque de la imagen \mathbf{x}^p variando λ de 0 a 1 con precisión 0,1, y seleccionamos el valor de λ , y restauración \mathbf{x}_λ^p , que minimizan $\|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}_\lambda^p\|^2$.

El procedimiento iterativo propuesto se resume en el Alg. 7. Usando la estimación inicial del bloque de la imagen $\mathbf{x}^p = \mathbf{y}_1^p$, el algoritmo alterna entre la estimación de los parámetros usando (5.33)-(5.37) y la estimación de los bloques de la imagen \mathbf{x}^p resolviendo (5.31) usando para ello el método de gradiente conjugado y evitando, con ello, el cálculo explícito de la matriz de covarianza $\Sigma_{\mathbf{x}}^p$.

Algoritmo 7: Método de deconvolución basado en combinación de modelos

Input: $\mathbf{y}_1^p, \mathbf{y}_2^p$ los bloques de las imágenes de tiempos largo y corto de exposición, respectivamente; y \mathbf{h}^p la PSF estimada en el parche p .

Output: \mathbf{x}^p el bloque restaurado.

for $\lambda = 0$ a 1 en pasos de 0,1 **do**

Set $\mathbf{x}_{new}^p = \mathbf{y}_1^p$

repeat

Set $\mathbf{x}_{old}^p = \mathbf{x}_{new}^p$

Calcular la estimación de los parámetros usando Ecs. (5.33)-(5.37) con $\mathbf{x}^p = \mathbf{x}_{old}^p$

Encontrar \mathbf{x}_{new}^p , la nueva estimación del bloque de imagen usando Ec. (5.31)

until $\frac{\|\mathbf{x}_{new}^p - \mathbf{x}_{old}^p\|^2}{\|\mathbf{x}_{old}^p\|^2} \leq threshold$

Set $\mathbf{x}_\lambda^p = \mathbf{x}_{new}^p$

end

Set $\mathbf{x}^p = \operatorname{argmin}_{\mathbf{x}_\lambda^p} \|\mathbf{y}_1^p - \mathbf{H}^p \mathbf{x}_\lambda^p\|^2$

Dotar de robustez a PSFs incorrectas

El método de deconvolución propuesto consigue estimar una imagen de alta calidad en casos en los que el bloque presenta un emborronamiento espacialmente invariante. Sin embargo, como ocurre con otros algoritmos de deconvolución, podría dar lugar a estimaciones erróneas de la imagen si la función de emborronamiento no se estima correctamente. Como ya mencionamos anteriormente, este problema se produce en bloques con poca textura y en parches en los que el emborronamiento varía dentro del bloque, como en zonas con objetos en movimiento u objetos a diferentes profundidades.

Para superar estos problemas, seleccionamos en cada bloque entre el resultado proporcionado por el Algoritmo 7 de deconvolución o el resultado de la eliminación de ruido del parche ruidoso \mathbf{y}_2^p . En este último caso, debido al corto tiempo de exposición, esperamos que la imagen \mathbf{y}_2^p no se vea afectada de emborronamiento. Como resultado, se espera una mejor estimación de la imagen nítida cuando eliminamos ruido en \mathbf{y}_2^p que aplicando el método de deconvolución en bloques en los que el núcleo no fue estimado con suficiente exactitud. Sin embargo, como se muestra en la sección de experimentos, la eliminación de ruido de toda la imagen no proporciona resultados suficientes, incluso aplicando el estado del arte en supresión de ruido [1]. Esto se debe principalmente a la pérdida de detalles en regiones con alta actividad espacial debido a la

corrupción por excesivo ruido.

En parches donde las estimaciones de la función de emborronamiento son erróneas, el parche tras la deconvolución diferirá mucho del bloque latente, además la estimación de β_2^p (la precisión del ruido en la imagen de corta exposición) generalmente resultará en un valor extremadamente pequeño en esos bloques. Podemos valernos de esta información para decidir entre aplicar nuestro algoritmo de deconvolución o el de eliminación de ruido. En concreto, proponemos calcular el valor medio de $\bar{\beta}_2^p$ de las estimaciones de β_2^p para todos los bloques, y aplicar eliminación de ruido en parches en los que se cumpla $t * \beta_2^p$ sea menor que el valor promedio, es decir

$$t * \beta_2^p < \bar{\beta}_2^p. \quad (5.42)$$

En nuestros experimentos encontramos que $t = 1,5$ es un valor adecuado. La eliminación de ruido se lleva a cabo por el método BM3D en [1]. Este método requiere la estimación de la varianza del ruido que se obtiene como el valor medio de β_2^{p-1} para todos los bloques en que se seleccionó nuestro método de deconvolución, es decir, no teniendo en cuenta los parches donde $t * \beta_2^p < \bar{\beta}_2^p$.

Cabe destacar, sin embargo, que una estimación incorrecta de la PSF, no necesariamente produce a una mala restauración, ya que el proceso de deconvolución tiene en cuenta tanto el bloque de exposición correcta como el de corta exposición. Pudiendo compensar una mala estimación de la función de emborronamiento, priorizando la información del parche de corta exposición. Sin embargo, los parches con un valor de β_2^p mucho menor que el valor medio de $\bar{\beta}_2^p$ son siempre malas restauraciones y por lo tanto, la sustitución de ellos por la versión del parche con supresión de ruido \mathbf{y}_2^p mejora la calidad de la restauración final.

5.6. Fusión de bloques

Después de procesar todos los bloques mediante los algoritmos descritos en las Secciones 5.4 o 5.5, las restauraciones obtenidas en cada uno de ellos van a ser fusionadas utilizando para ello una función de enventanado normalizada \mathbf{s} , como la que mostramos en la Figura 5.1, que evita artefactos entre bloques. La aplicación del enventanado se realiza multiplicando cada uno de los bloques por la función de enventanado que ha sido previamente normalizada de forma que la suma de todos los factores para un determinado píxel sea igual a 1, es decir,

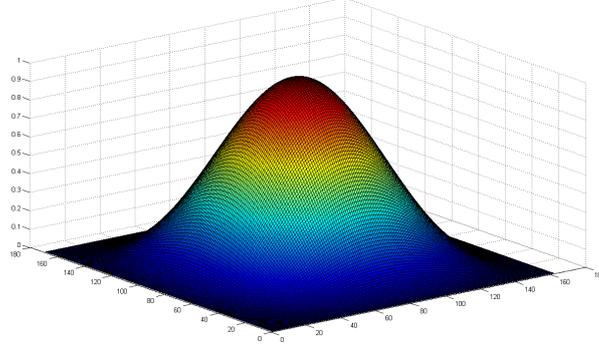


Figura 5.1: Ventana de Hann utilizada en la etapa de fusión de bloques.

$$\mathbf{x} = \sum_p^P \mathbf{s}^p \mathbf{x}^p, \quad (5.43)$$

$$\sum_p^P \mathbf{s}_l^p = 1, \text{ para } 1 \leq l \leq (N_x \times N_y), \quad (5.44)$$

con P el número de parches. Podemos utilizar varias funciones de enventanado tales como: Gaussiana, rectangular, Hamming o de Hann. Sin embargo, el tipo de ventana de Hann, definida en unidimensional como $s_b^p(j) = 0,5 \left(1 - \cos \left(\frac{2\pi j}{b-1} \right) \right)$, con $b = B_x, B_y$, produce sensiblemente mejores resultados. Nótese que la extensión al caso bidimensional es directo realizando el producto exterior de los vectores, $\mathbf{s}_{B_y}^p$ y $\mathbf{s}_{B_x}^p$ de tamaño B_y, B_x , respectivamente, obteniendo $\mathbf{s}^p = \mathbf{s}_{B_y}^p \times (\mathbf{s}_{B_x}^p)^t$, una ventana de tamaño $(B_y \times B_x)$ para cada uno de los parches p . Este tipo de enventanado además, ha sido utilizado y recomendado en otros trabajos como [84].

5.7. Experimentos

En esta sección, se analiza la eficacia de los algoritmos desarrollados en pares de imágenes sintéticas y reales adquiridas en entornos de escasa luminosidad con diferentes cámaras digitales. Con el fin de generar un conjunto de datos sintéticos, dividimos las imágenes en parches sin solapamiento y aplicamos una función de emborronamiento local invariante para cada bloque, lo que genera un emborronamiento globalmente variante.

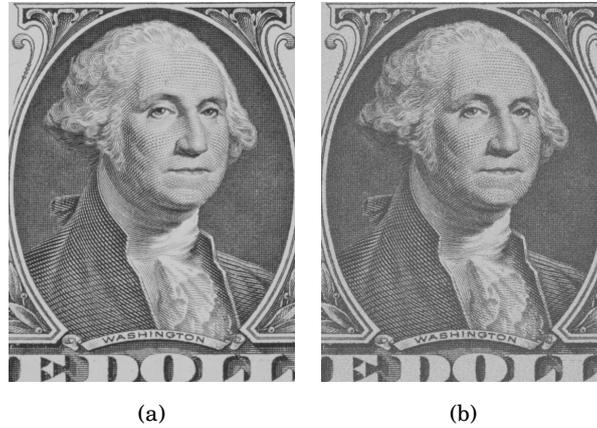


Figura 5.2: (a) Imagen original, (b) imagen subexpuesta simulada, calibrada fotométricamente con la imagen borrosa mostrada en la Figura 5.3c.

5.7.1. Experimentos sintéticos

El objetivo de este experimento sintético es, suponiendo un núcleo invariante localmente pero globalmente variante, evaluar tanto la estimación del núcleo como la precisión del algoritmo de deconvolución. Téngase en cuenta que el procedimiento de deconvolución explota la información de ambas observaciones: la imagen borrosa y la de exposición corta.

El conjunto de imágenes borrosas simulando un tiempo de exposición largo se obtienen de la siguiente forma: Primeramente la imagen original, mostrada en la Figura 5.2a, de tamaño 800×600 , se divide en 4×3 bloques sin solapamiento. Utilizando cuatro emborronamientos espacialmente variantes distintos generamos las cuatro imágenes borrosas que se muestran en la fila superior de la Figura 5.3. Las distintas PSFs con la que convolucionamos la imagen original se muestran en la segunda fila de la Figura 5.3. Cada función de emborronamiento (individual) es de tamaño 15×15 .

Con el fin de simular una imagen subexpuesta realista, empleamos el proceso de adquisición de imágenes descrito en [12] y simulamos la corrupción por ruido en el dominio de irradiancia. En primer lugar, la imagen original en la Figura 5.2a se convirtió al dominio de irradiancia des-haciendo la corrección gamma, suponiendo $\gamma = 2,2$. Posteriormente, asumiendo que la imagen de corta exposición posee un valor de exposición (EV) seis pasos inferior a la imagen original, dividimos cada valor de pixel por $2^6 = 64$. Añadimos ruido gaussiano de media cero y varianza 199,3 correspondiente a una SNR de 10 dB. Finalmente, a la imagen resultante se le aplicó una

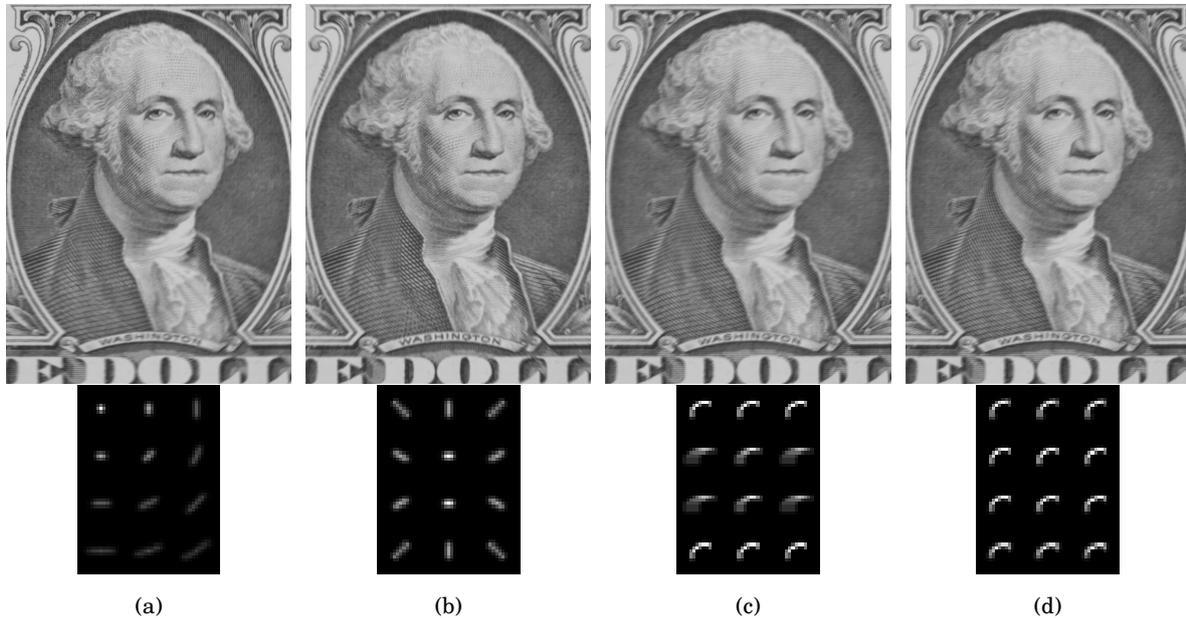


Figura 5.3: Imágenes borrosas simulando larga exposición. Las PSFs utilizadas para generar cada una de las imágenes borrosas se muestran en la segunda fila. Nótese que los valores de las funciones de emborronamiento se han llevado al rango $[0,255]$ por razones de visualización.

corrección gamma con $\gamma = 2,2$ y se cuantificó a ocho bits para obtener la imagen subexpuesta simulada.

Las imágenes fueron calibradas fotométricamente usando ecualización de histograma con el histograma de la imagen borrosa como describimos en el Capítulo 3. En concreto, la observación ruidosa calibrada de la Figura 5.2b se obtiene a partir de la ecualización del histograma con la imagen borrosa de la Figura 5.3c.

Los métodos de deconvolución propuestos en este capítulo (resumidos en los Algoritmos 5 y 7) han sido ejecutados utilizando un valor de $threshold = 10^{-5}$ para detener el proceso iterativo. Los núcleos de emborronamiento estimados se representan en la segunda fila de las Figuras 5.4 y 5.5 para los Algoritmos 3 y 6, respectivamente. Nótese que las PSFs obtenidas son exactamente las mismas para ambos algoritmos ya que, al no tener solapamiento en este experimento sintético, el paso de corrección de la PSF descrito en la Sección 5.4.1 no produce cambios en la estimación de la PSF. Sin embargo, más abajo diseñamos otro experimento sintético con solapamiento en el que se corrigen las PSFs.

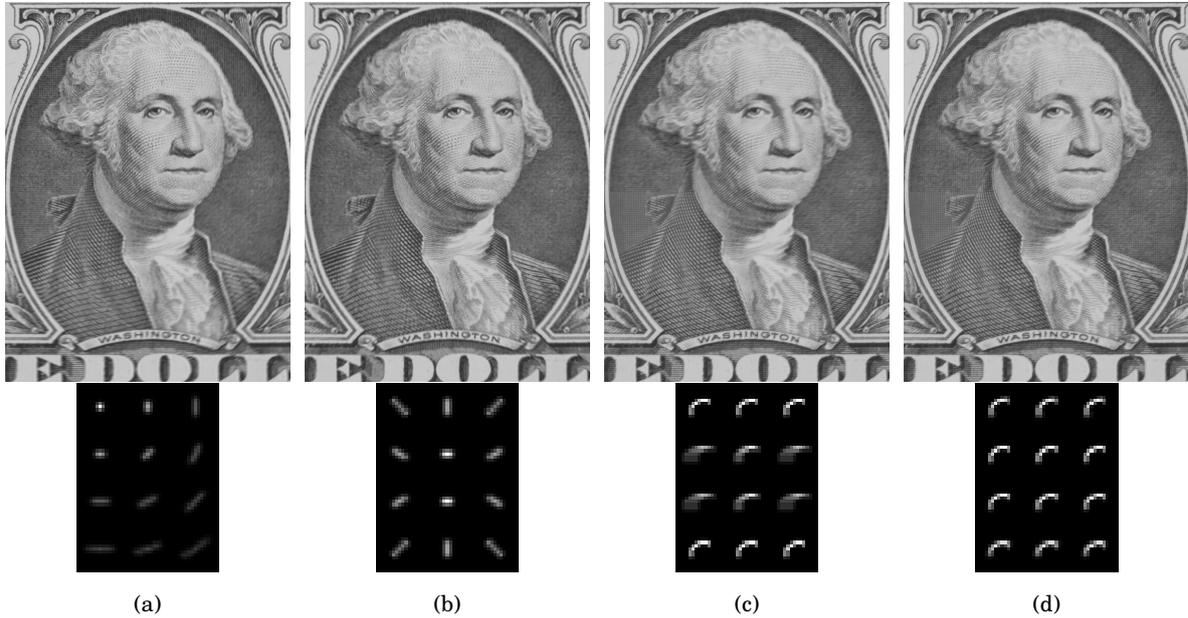


Figura 5.4: Restauraciones resultado de aplicar el Algoritmo 3 (AREVDR) en las imágenes sintéticas de la Figura 5.3. En la fila inferior se muestran los núcleos de emborronamiento estimados.

Es evidente que la etapa de estimación del núcleo de la Sección 5.3 recupera de manera fidedigna cada una de las PSFs, hecho que era previsible; ya que cada bloque fue emborronado con un único núcleo de emborronamiento y, por lo tanto, ninguno de los parches restaurados fueron reemplazados por el resultado de la eliminación de ruido en ninguno de los cuatro experimentos sintéticos.

Los resultados de aplicar el método en 5 pueden observarse en la Figura 5.4; este método consigue eliminar el emborronamiento más importante en un tiempo muy ajustado, téngase en cuenta que el tiempo empleado en el proceso de deconvolución descrito en la Sección 5 no supuso más de 4 segundos en ninguno de los experimentos sintéticos; mientras que el tiempo consumido en la estimación de los núcleos de emborronamiento fue de más de 50 segundos. La Tabla 5.4 muestra los tiempos de ejecución para el experimento de la Figura 5.3c. Los algoritmos se ejecutaron en un ordenador portátil Core 2 Duo a 2,13 GHz. Téngase en cuenta que la implementación actual de los métodos propuestos, son implementaciones en MATLAB no optimizadas para su uso en máquinas masivamente paralelas. Nótese que el tiempo total del AREVDCM fue aproximadamente 6 veces mayor que el AREVDR.

Cuadro 5.1: Métricas PSNR y SSIM (Peak Signal to Noise Ratio y Structural Similarity) para las imágenes de los experimento sintéticos de la Figura 5.3.

	Conjunto de núcleos de emborronamiento							
	a		b		c		d	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Borrosa	23.88	0.75	23.79	0.75	20.99	0.51	20.80	0.49
Ruidosa	22.80	0.85	22.80	0.85	22.71	0.85	22.77	0.85
Denoising [1]	24.41	0.89	24.40	0.89	24.07	0.89	24.25	0.89
deconvblind (esp-var.)	22.95	0.87	21.87	0.88	22.07	0.85	22.10	0.89
deconvblind (esp-invar.)	22.30	0.78	22.06	0.78	18.76	0.80	18.57	0.83
Mét. en [2] (esp-var.)	22.69	0.86	20.55	0.85	22.53	0.87	23.31	0.90
Mét. en [2] (esp-invar.)	22.09	0.77	21.33	0.76	22.43	0.85	23.47	0.89
AREI	22.61	0.78	24.12	0.84	25.88	0.83	26.60	0.86
AREVDR	25.88	0.85	26.38	0.87	27.45	0.89	28.25	0.92
AREVDCM ($\lambda = 0$)	25.46	0.85	25.48	0.87	20.51	0.58	20.19	0.56
AREVDCM ($\lambda = 1$)	27.77	0.94	28.49	0.94	27.45	0.93	28.31	0.94
AREVDCM (λ estimado)	28.26	0.95	29.15	0.95	28.26	0.95	28.74	0.95
AREVDCM (λ óptimo)	28.62	0.95	29.53	0.96	29.08	0.96	30.74	0.97

Para obtener las mejores restauraciones aplicando AREVDCM (Alg. 6), se ha variado el valor λ , en el rango [0,1] con precisión de 0,1 y tomado el valor óptimo (en PSNR) para cada uno de los bloques. Las restauraciones obtenidas siguiendo este procedimiento se muestran en la primera fila de la Figura 5.5. Por otro lado, los resultados usando valores estimados de λ , se muestran en la segunda fila de la Figura 5.5. Examinando ambos resultados es difícil encontrar diferencias excepto que las restauraciones obtenidas con los valores óptimos de λ parecen ser ligeramente más nítidas. Los valores óptimos y nuestras estimaciones de λ para cada uno de los bloques de la Figura 5.5c aparecen en la Tabla 5.3. Como se desprende de la tabla, tanto los valores estimados y óptimos son bastantes parecidos, lo que confirma que el método de

Cuadro 5.2: Comparación en PSNR para el conjunto de PSFs estimadas en los experimentos sintéticos.

	Conjunto de emborronamientos			
	a	b	c	d
deconvblind (esp-var.) (inicialización estándar)	84.12	84.20	84.67	84.00
deconvblind (esp-var.) (inicializada con nuestra PSF)	96.58	96.19	93.62	92.83
Método en [2] (esp-var.)	90.90	91.63	90.03	89.18
Métodos propuestos (Secs. 5.4 y 5.5)	96.40	96.14	93.63	91.49

Cuadro 5.3: Valor de λ óptimo a la izquierda y el estimado a la derecha para cada uno de los bloques de la Figura 5.3c. Para la ejecución de AREVDCM (Alg. 6)

0.6	0.7	0.6	0.5	0.7	0.5
0.8	0.7	0.8	0.8	0.7	0.8
0.6	0.9	0.8	0.5	0.8	0.8
0.6	0.6	0.5	0.5	0.6	0.4

estimación propuesto para el valor de λ se ajusta fielmente a lo esperado.

Los algoritmos propuestos se han comparado con varios métodos de restauración de imágenes tales como, el método de eliminación de ruido en [1], el algoritmo de deconvolución ciega basado en Richardson-Lucy (RL), implementado por la rutina de MATLAB `deconvblind`, el estado del arte en exposición dual descrito en el trabajo [2] y el algoritmo de restauración espacialmente invariante, (AREI), propuesto en el Capítulo 4. Nótese que todos estos métodos de deconvolución asumen un emborronamiento espacialmente invariante. Para llevar a cabo una comparación justa, también se aplican los algoritmos anteriores utilizando el marco espacial-

Cuadro 5.4: Tiempo de ejecución (en segundos) de cada uno de los métodos comparados para los experimentos en la imagen de la Figura 5.3c.

Denoising [1]	deconvblind (esp-var.)	Mét. en [2]	AREI (Cap. 4)	AREVDDR	AREVDCM
6.9	13.5	380	390	54.6	288

mente variante por bloques, descrito en el Alg. 6 realizando el paso de deconvolución con cada uno de los métodos que comparamos. En este proceso se obtiene una restauración para cada bloque, en lugar de una única global para lo que fueron diseñados originalmente.

En la Tabla 5.1 se muestran las métricas PSNR y SSIM para las imágenes restauradas obtenidas por los métodos propuestos y los algoritmos competidores. El conjunto de núcleos de emborronamiento variante también se evaluaron numéricamente usando la métrica PSNR y los resultados se muestran en la Tabla 5.2. Nótese que únicamente se compararon (en PSNR), las PSFs estimadas en un marco espacialmente variante, que fueron como se generaron originalmente las observaciones borrosas. Para poder realizar una comparación visual de las restauraciones, presentamos los resultados de la aplicación de todos los métodos al par de imágenes de las Figuras 5.3c y 5.2b. Las restauraciones se presentan en la Figura 5.6 y las estimaciones de los núcleos en la Figura 5.7. Claramente, la eliminación de ruido de la imagen ruidosa no parece suficiente, debido al suavizado de los detalles. Por otro lado, también ejecutamos el algoritmo de deconvolución de RL en el que especificamos una matriz de unos, a la función `deconvblind`, pero esta inicialización de la PSF obtuvo estimaciones poco fiables, por ello, para realizar comparaciones más justas preferimos inicializar `deconvblind` usando las PSFs estimadas en [2] (para el caso espacialmente invariante) o los núcleos estimados siguiendo nuestro procedimiento descrito en la Sección 5.3 (para el caso espacialmente variante). Los resultados pueden consultarse en la Tabla 5.2. En ambos casos el número de iteraciones se fija a 20. Los resultados confirmaron que la rutina `deconvblind` no consigue mejorar sustancialmente (en PSNR) la función de emborronamiento. Los resultados en este caso se presentan en las Figuras 5.6b y 5.6c para las ejecuciones del algoritmo RL espacialmente variante e invariante, respectivamente. Ambas restauraciones exhiben un efecto anillado considerable, más perceptible al acercarnos a los bordes de la imagen y especialmente notable cuando lo aplicamos en nuestro marco espacialmente variante. Por otra parte, el método propuesto en [2] proporciona métricas muy similares cuando se aplica en un enfoque espacialmente variante, aunque con resultados visuales algo mejores (véase la Figura 5.6d y la métrica SSIM en la Tabla 5.1). Las diferencias en los resultados de aplicar [2] (en su versión espacialmente variante) y los métodos propuestos en éste capítulo, se deben al tamaño de los bloques; ya que la estimación del emborronamiento que se propone en la versión espacialmente variante del método en [2], precisa de un mayor número de píxeles para obtener una estimación de emborronamiento más acertada. La aplicación de la versión original del método [2] con estimación de parámetros automática,

no consigue eliminar parte del emborronamiento de algunas zonas como demuestra la Figura 5.6e, véase como se aprecia efectos dientes de sierra y de anillado cerca de las fronteras. Por otro lado, el método propuesto en el Capítulo 4, mejora a los métodos que consideran un emborronamiento espacialmente invariante pero con resultados algo alejados de los métodos que proponemos en este capítulo. En este caso, hemos preferido ejecutar la versión original del método propuesto en el Capítulo 4, sin considerar la versión espacialmente variante del método por varias razones: para hacer comparaciones justas habría que estimar el parámetro λ y además la ejecución del método en un marco espacialmente variante adolece del mismo inconveniente que la versión espacialmente variante de [2], es decir al dividir las imágenes de entrada en bloques pequeños, no tenemos información suficiente para realizar una estimación fiable y estable del emborronamiento y la imagen conjuntamente.

El método propuesto con $\lambda = 0$ (es decir, usando únicamente el modelo de imagen SAR) produce restauraciones suaves, como puede observarse en la Figura 5.6h. Mientras que por el contrario, cuando se usa exclusivamente el modelo TV como modelo de imagen (estableciendo $\lambda = 1$), los bordes son más nítidos, como se aprecia en la estimación de la imagen de la Figura 5.6i pero presenta artificios en zonas planas debido a la presencia de ruido. Cuando ajustamos el parámetro λ , y lo ajustamos al valor óptimo para cada bloque, se obtienen mejores resultados en comparación con el uso de un único modelo de imagen (véase, por ejemplo, la textura de la chaqueta de G. Washington de la Figura 5.6j).

Por último, se analizó el efecto de la etapa de corrección del AREVDR (Alg. 3) y del umbral t (Ec. 5.42) en el AREVDCM (Alg. 6). Para ello, ejecutamos el método propuesto en los 4 experimentos de las Figuras 5.3 y 5.2b pero en este caso, considerando 7×5 bloques con un solapamiento del 50%. Esto implica que 4×3 bloques presentarán un único emborronamiento dentro del bloque, ya que coinciden con los bloques sin superposición en la imagen. El resto de bloques tendrán partes afectadas por emborronamientos diferentes. La Figura 5.8d muestra el resultado de aplicar el AREVDR en el experimento de la Figura 5.3b considerando la configuración de 7×5 bloques solapados. La etapa de corrección se aplica en los bloques de la PSF de la Figura 5.8a que se muestran en color blanco en la Figura 5.8b. Los beneficios de esta etapa van a ser mucho mayores en zonas donde el núcleo no pudo estimarse correctamente ya que podría recuperarse de los bloques adyacentes. La Figura 5.8c muestra el efecto de la reparación, es decir, en algunos bloques es ventajoso aplicar la corrección (suavizado de ruido o recuperación de la PSF) mientras en otros no lo es tanto (podría reducir las dimensiones de éste). La PSNR

obtenida en este experimento sintético fue de 25,6 dB. Los resultados podrían mejorarse haciendo uso de los bloques con eliminación de ruido, desafortunadamente, en este experimento concreto ningún bloque fue substituido por su correspondiente con eliminación de ruido, debido a que la etapa de corrección ha promediado un gran número de PSFs (Figura 5.8b).

A continuación mostramos el efecto del umbral t en el AREVDCM. En este experimento demostramos que la información de la imagen de corta exposición puede ser muy útil cuando el proceso de deconvolución no funciona apropiadamente para un bloque en particular. Esto ocurre con frecuencia en los bloques que presentan varios emborronamientos dentro de un mismo bloque provocando una estimación errónea de la PSF. La Figura 5.9a presenta los valores de PSNR variando el umbral t desde 1 hasta 2 en pasos de 0,1 para el experimento de la Figura 5.3b (con la configuración de 7×5 bloques solapados). Para proporcionar una confianza ligeramente superior a los bloques resultado de la deconvolución fijamos $t = 1,5$. En la Figura 5.9d puede observarse la imagen resultado para el experimento de la Figura 5.3b con un valor de umbral $t = 1,5$, mientras que la Figura 5.9b muestra las estimaciones de la PSF. Téngase en cuenta que los cuatro bloques substituidos por su versión sin ruido, (Fig. 5.9c), corresponden a bloques afectados por diferentes emborronamientos.

5.7.2. Experimentos reales

En la siguiente serie de experimentos, aplicamos los métodos propuestos a tres pares de imágenes reales adquiridas en entornos de baja iluminación con diferentes cámaras digitales. Los pares de imágenes se muestran en la Figura 5.10. Las imágenes de entrada han sido calibradas fotométricamente y geométricamente tal y como se describe en el Capítulo 3. Nótese que, en nuestros algoritmos, la ecualización de histograma se realiza exclusivamente en la banda de luminancia; sin embargo, para mostrar las Figuras 5.10b, 5.10d y 5.10f, las bandas de crominancia de la imagen de corta exposición fueron también calibradas fotométricamente usando la técnica de ecualización de histograma. En todos los experimentos se utilizó un porcentaje de solapamiento entre bloques del 50%, con una ventana de Hann para la etapa de post-procesamiento, parando el algoritmo cuando se obtuvo un valor de convergencia en iteraciones consecutivas menor de $threshold = 10^{-5}$. Nótese que en cada uno de los experimentos reales, se utilizó la misma configuración de bloques para ejecutar los dos algoritmos propuestos, y han sido comparados con el estado del arte en exposición dual propuestos en [3] y [4], además comparamos con el estado del arte en algoritmos de supresión de ruido basado en empareja-

miento de bloques en [1], usando, en este caso, únicamente la imagen de corta exposición.

El primer experimento se llevó a cabo sobre el par de imágenes, de tamaño 512×512 y mostradas en las Figuras 5.10a y 5.10b, usadas por primera vez en el trabajo [42]. Las imágenes fueron captadas con la cámara Canon EOS 400D. El experimento exhibe una escena con emborronamiento provocado por objetos en movimiento y un ligero movimiento de la cámara. El tiempo de exposición fue de $1/4$ s para la captación de la imagen borrosa y $1/15$ s, con una compensación de exposición (EV) de -2, para la imagen de exposición menor. Ambos disparos se efectuaron usando un factor ISO 400. Dado que el emborronamiento no parece severo, se utilizó un tamaño para cada PSF de 15×15 píxeles, con un total de 6×6 parches. La PSF estimada por el método [2], en este experimento, se muestra en la Figura 5.14a y la PSF estimada por el método descrito en el Cap. 4 (AREI) se muestra en la Figura 5.14d. Ninguna de las PSFs anteriores consigue modelar correctamente el movimiento de los objetos de la escena y por tanto las restauraciones no van a poder deshacer las degradaciones sufridas. Nótese que el núcleo obtenido por el método AREI se aproxima a la función delta. El núcleo de emborronamiento estimado y corregido en cada uno de los bloques para el AREVDR, se representa en las Figuras 5.15a y 5.15c, respectivamente. De forma análoga, las PSFs estimadas por el AREVDCM, se representa en la Figura 5.16a. Los bloques que no fueron restaurados adecuadamente por el AREVDR (Alg. 3), se representan en color blanco en la Figura 5.15d, mientras que los bloques de la estrategia de selección del AREVDCM (Alg. 6) se muestran en la Figura 5.16b. Según los algoritmos propuestos, estos parches se sustituyeron por el bloque correspondiente de la imagen con supresión de ruido siguiendo el algoritmo en [1]. Aunque, claramente, la etapa de corrección del emborronamiento del AREVDR consigue recuperar parte de las PSFs erróneas, no parece suficiente para evitar el uso del método de eliminación de ruido en esos bloques.

Nótese como los dos algoritmos propuestos en este capítulo, resuelven satisfactoriamente el problema de emborronamiento por objetos en movimiento. En el primero de ellos, se aprecia un mayor suavizado en la restauración (Fig. 5.11e) debido a un mayor uso de la eliminación de ruido mientras que el segundo de los métodos propuestos solucionó correctamente algunos parches con objetos en movimiento, como el niño en el lado derecho de la imagen en la Figura 5.10a, ya que en estos casos el algoritmo da mayor prioridad a la información de la observación ruidosa. La imagen resultante, que se muestra en la Figura 5.11f, presenta un mayor contraste y menor ruido en el fondo que la imagen con eliminación de ruido de la Figura 5.11a. A su vez la restauración parece más realista que la obtenida con los métodos que consideran emborrona-

miento espacialmente invariante [2] y el AREI del Capítulo 4 (véase, la persona que acompaña al niño en bicicleta en el margen izquierdo de la imagen se encuentra borrosa debido al emborronamiento espacialmente variante). Los resultados por los métodos aportados en el presente capítulo son ligeramente mejores que los ofrecidos por el método [3], representado en la Figura 5.11c. Véase como el árbol del fondo de la imagen exhibe más detalles en la restauración obtenida por cualquiera de los algoritmos propuestos.

El segundo experimento se llevó a cabo sobre el par de imágenes de tamaño 512×512 de las Figuras 5.10c y 5.10d. En este experimento evaluamos el rendimiento del algoritmo en situaciones con un emborronamiento por movimiento de cámara severo con objetos a distintas profundidades junto a una imagen de corta exposición con un nivel de ruido elevado. Las imágenes fueron captadas con la cámara Olympus C5060WZ. El tiempo de exposición de la imagen borrosa es 1,3s con factor ISO 100, siendo el factor ISO de 400 y 1/100s de tiempo de exposición para la imagen de exposición más corta. El número de bloques se establece a 6×6 , con un tamaño de PSF de 51×51 . El ruido en la imagen de corta exposición después del calibrado es tan severo que el método de eliminación de ruido en [1], sólo permite recuperar algunos detalles en torno a la taza donde el ruido es menos acusado (como se aprecia en la Figura 5.12a). Tanto el método en [2] (Figura 5.12b) como el AREI propuesto en el Capítulo 4 (Figura 5.12d), no proporcionaron estimaciones adecuadas del núcleo de emborronamiento (véase Figuras 5.14b y 5.14e) debido a la naturaleza espacialmente variante del emborronamiento por desplazamiento y giro de la cámara y por consiguiente la imagen restaurada en este caso aparece suavizada en exceso. Por otro lado, el método en [3], consigue recuperar detalles de la imagen borrosa tales como algunas zonas próximas a la taza y al jarrón (obsérvese la Figura 5.12c); sin embargo, tanto el jarrón como la taza aparecen también suavizados con algunos contornos artificiales y otros que fueron eliminados. En concreto, la parte inferior de la vasija no posee ninguno de los reflejos de la taza que si son visibles en las restauraciones con los métodos propuestos en el presente capítulo (véase las Figuras 5.12e y 5.12f). También pueden observarse en estos mismos resultados, como recuperamos parte del título en el tomo de los libros, en parte, gracias a la correcta estimación de la rotación producida por el movimiento de la cámara y que varía a lo largo de la imagen mientras se realizó la captación (véase la Figura 5.16c). Los algoritmos propuestos sugieren la utilización de los bloques con eliminación de ruido en zonas alrededor de la taza (Figs. 5.16d y 5.15h) en lugar de los obtenidos del proceso de deconvolución. Nótese como ahora son visibles los pequeños detalles como las letras en la taza o en el libro detrás de

la vasija. Aunque visiblemente más nítidos en la restauración que proporciona el AREVDCM (Figura 5.12f), Por otro lado véase como la etapa de corrección del AREVDR selecciona los bloques en los que efectivamente no pudo recuperarse un núcleo correcto (Fig. 5.15f) que coinciden con la parte del jarrón.

En el último experimento se utilizaron el par de imágenes de las Figuras 5.10e y 5.10f de tamaño 800×600 . La imagen de exposición correcta presenta dos tipos de emborronamientos: el primero de ellos debido al movimiento de la cámara y un segundo emborronamiento producido por una reducida profundidad de campo (DOF, *depth of field*). Las imágenes se tomaron con una cámara Canon EOS 500D. El tiempo de exposición de la imagen borrosa es $1/10$ s mientras que en la imagen de menor exposición se utilizó $1/160$ s con una compensación de exposición (EV) de -4, el factor ISO en ambas tomas se fijó a 100. La imagen presenta un emborronamiento considerable, por ello se utilizó un tamaño de PSF de 31×31 píxeles y 8×6 parches. La imagen de menor exposición de la Figura 5.10f contiene ruido de cuantificación que no se consiguió eliminar por el método de supresión de ruido en [1], como se muestra en la chimenea de la parte superior de la Figura 5.13a, tampoco consigue suavizar el ruido de la reja en la parte inferior izquierda de la imagen. La restauración mediante el AREVDR obtiene un resultado muy parecido al propuesto en [4] mostrado en la Figura 5.13c; pero la restauración notablemente más nítida que los competidores es la producida por el AREVDCM (Fig. 5.13f). El método [2] (Fig. 5.13b), al igual que el resultado del método AREI de la Figura 5.11d, no parecen efectivos en este experimento debido precisamente al emborronamiento espacialmente variante que exhibe la imagen borrosa, las PSFs obtenidas se representan en las Figuras 5.14c y 5.14f, en ambas estimaciones, se realiza un promedio de los emborronamientos a distintas profundidades, es por ello que las restauraciones deshacen parcialmente el emborronamiento del fondo (véase el tejado de la casa). Sin embargo, los métodos propuestos en el presente capítulo resuelven el problema de reducción de la profundidad de campo y reemplazan la farola de la parte más próxima al observador por la versión sin ruido que, además, presenta un nivel de saturación menor y, por tanto, una mejor calidad. Obsérvese que no fue posible obtener una estimación de la PSF correcta en los bloques alrededor de las farolas debido a la saturación de los píxeles, tampoco en la zona con árboles y la pared del fondo claramente con emborronamientos distintos debido a la diferencia de profundidades. En este aspecto, la etapa de corrección del AREVDR realiza un buen trabajo, reparando algunos bloques con estimaciones de núcleo incorrectas (Fig. 5.15j) y evitando un uso abusivo de la eliminación de ruido (5.15l). Por otro lado la estrategia del AREVDCM, prefiere

mantener esas estimaciones de emborronamiento incorrectas pero sustituirlas por los bloques con supresión de ruido en esas áreas (consultar Figura 5.16f).

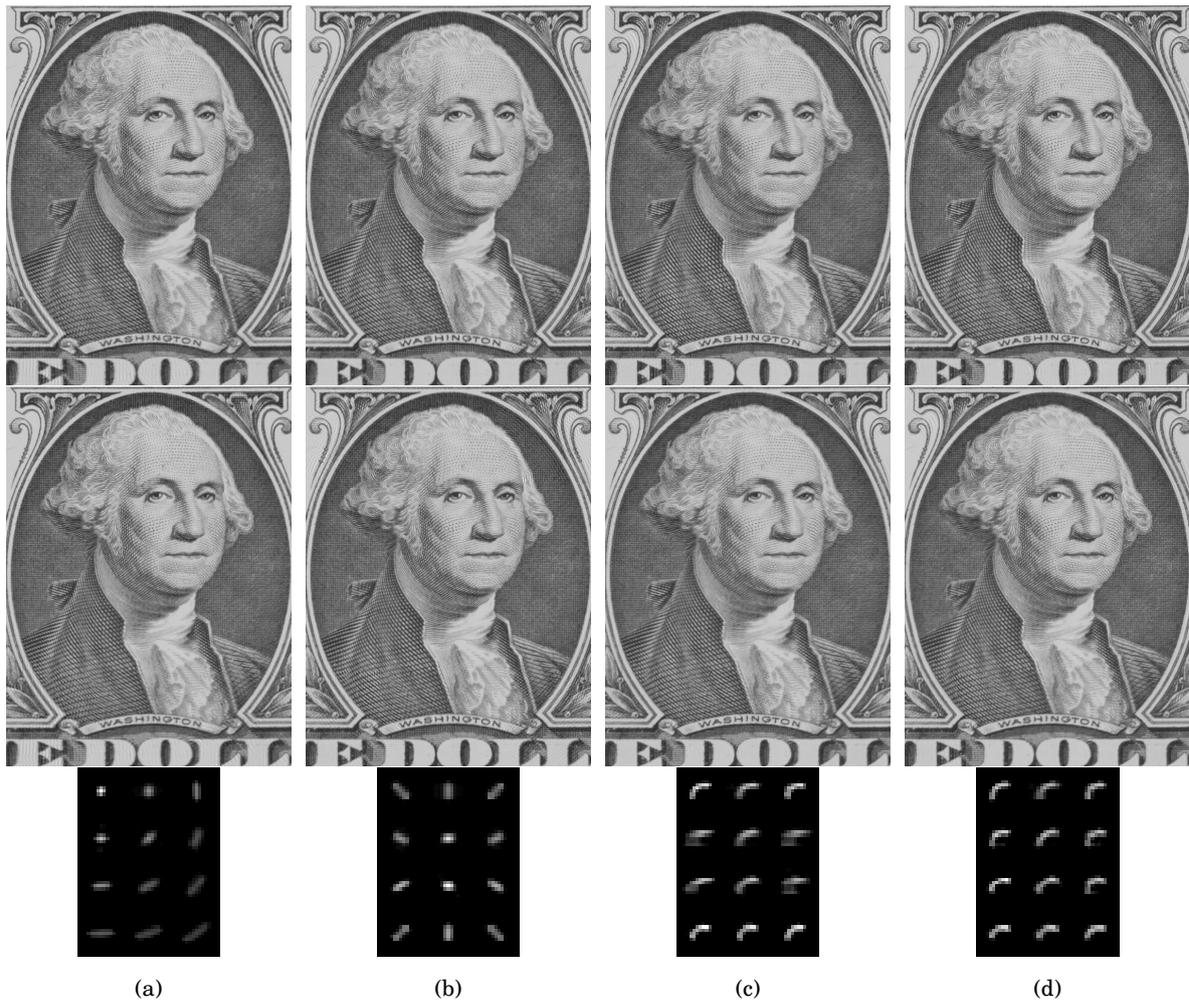


Figura 5.5: Restauraciones resultado de aplicar el Algoritmo 6 (AREVDCM) en las imágenes sintéticas de la Figura 5.3. Fila superior: Usando el valor óptimo de λ . Fila central: Usando el valor estimado de λ . Fila inferior: Núcleos de emborronamiento estimados.



Figura 5.6: Restauraciones resultado del experimento sintético de la Figura 5.3c y 5.2b: (a) imagen con supresión de ruido usando el método [1], (b) *deconvblind* aplicando nuestro marco de bloques para emborronamiento espacialmente variante, (c) *deconvblind* asumiendo emborronamiento espacialmente invariante, (d) método en [2] (esp-var.), (e) método original en [2], (f) AREI, (g) AREVDR, (h) AREVDCM ($\lambda = 0$), (i) AREVDCM ($\lambda = 1$), (j) AREVDCM.

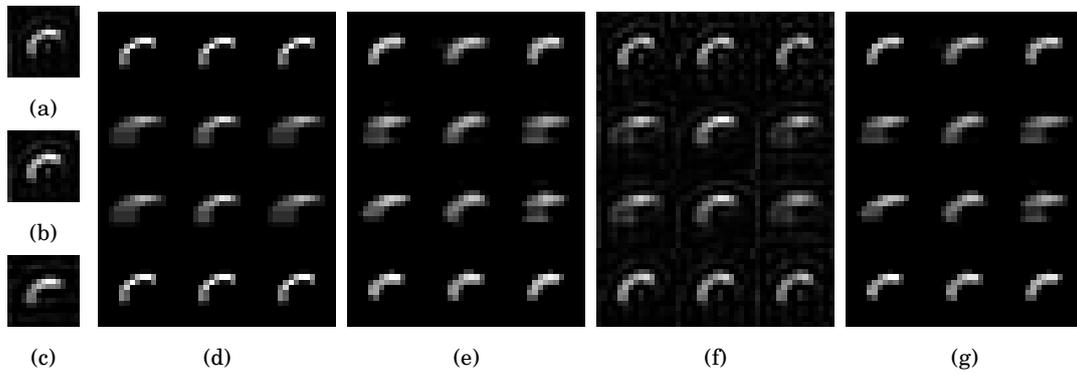


Figura 5.7: PSFs estimados en los experimentos de la Figura 5.6: (a) PSF global estimada en `deconvblind`, (b) PSF global en [2], (c) PSF global en AREI, (d) PSFs reales, (e) PSF espacialmente variante estimada en `deconvblind`, (f) PSFs asumiendo emborronamiento espacialmente variante en [2], (g) PSFs estimadas en los métodos propuestos. Cada PSF individual de tamaño 15×15 .

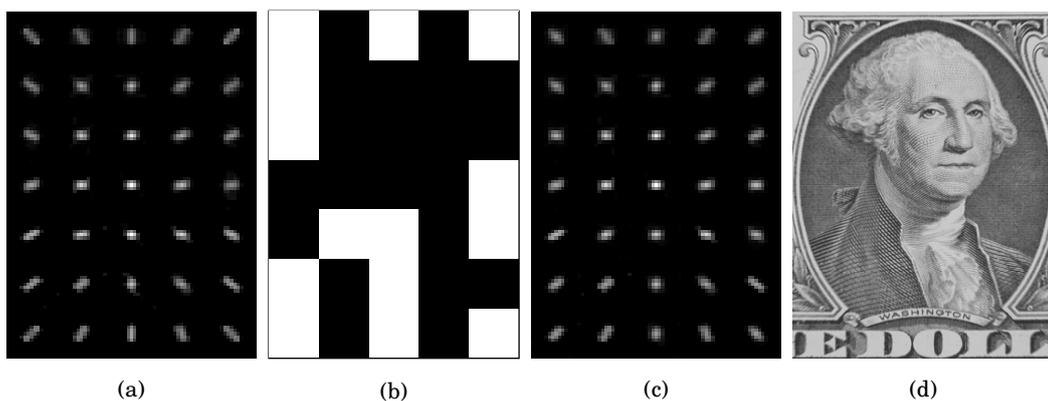


Figura 5.8: Resultados del experimento sintético de la Figura 5.3b con solapamiento y configuración 7×5 bloques aplicado a AREVDR. (a) PSFs estimadas. (b) bloques donde se ha aplicado la etapa de corrección de emborronamiento. (c) PSFs corregidas. (d) Imagen restaurada con una PSNR de 25,6 dB.

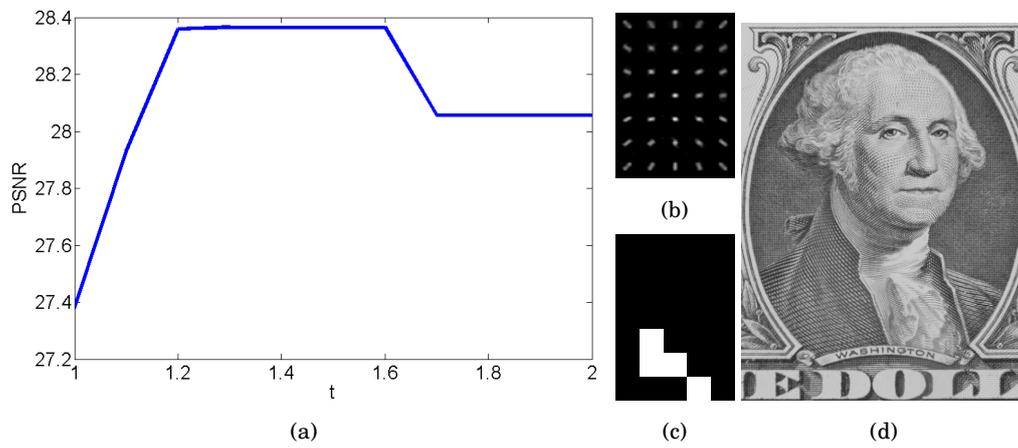


Figura 5.9: Resultados del experimento sintético de la Figura 5.3b con solapamiento y configuración 7×5 bloques aplicado a AREVDCM: (a) Evolución de la PSNR para distintos valores del umbral t usando los valores estimados de λ . (b) PSF estimadas. (c) bloques con eliminación de ruido con $t = 1, 5$. (d) Imagen restaurada con $t = 1, 5$.

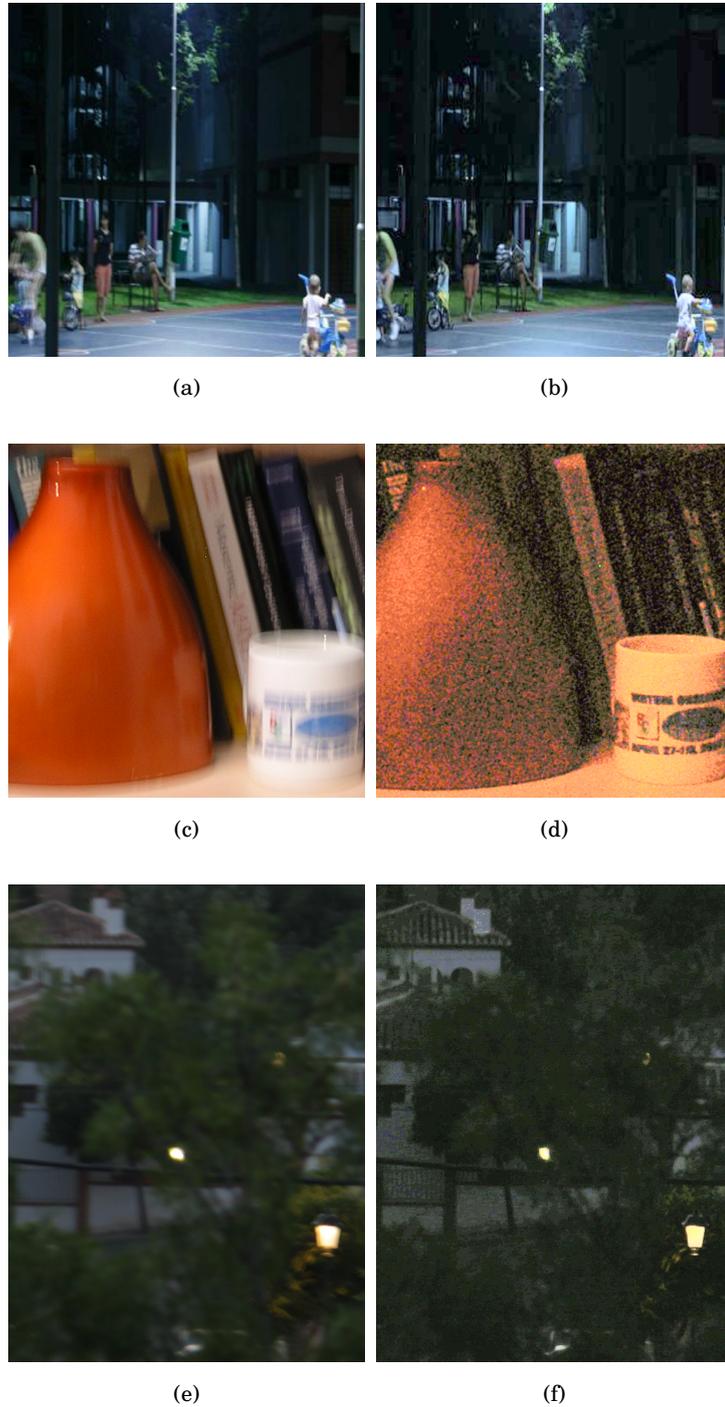


Figura 5.10: Imágenes de prueba después del calibrado: Izquierda, imagen de larga exposición y derecha, imagen de corta exposición observada.

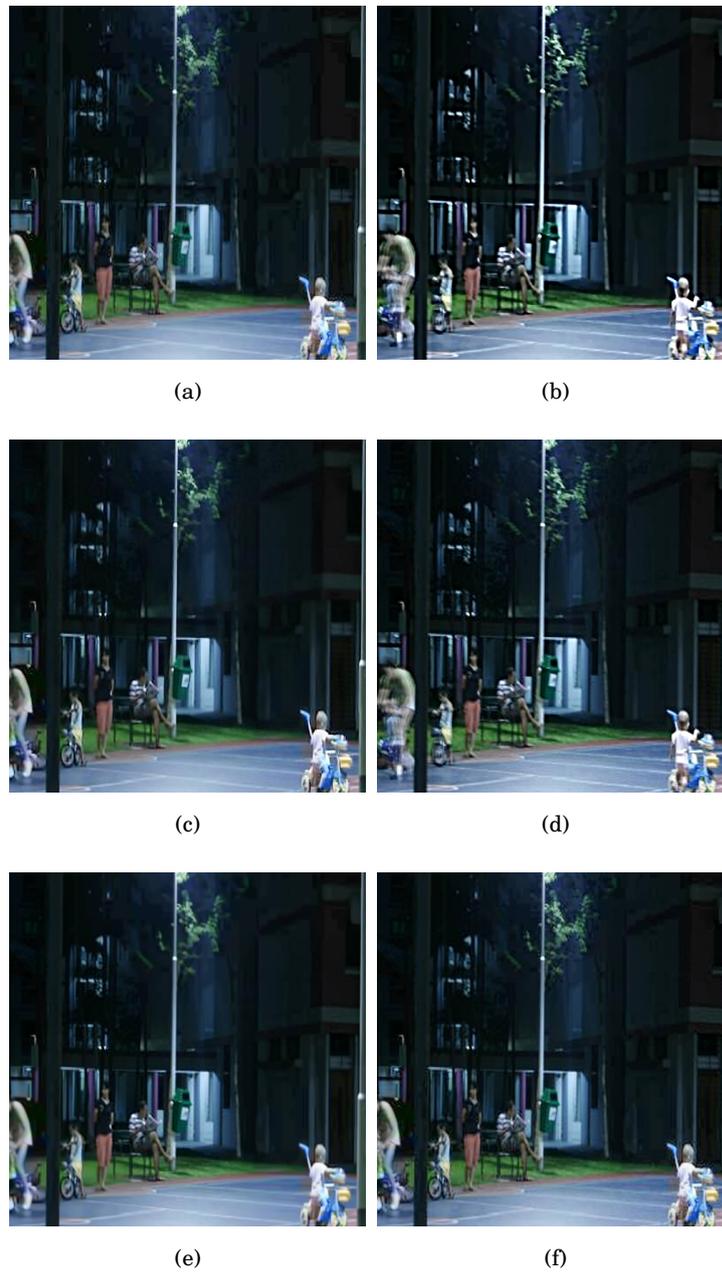


Figura 5.11: Resultados para el primer experimento: (a) imagen tras supresión de ruido usando [1], (b) restauración con el método en [2] (c) imagen restaurada con el método [3],(d) resultado de aplicar AREI, (e) resultado de aplicar AREVDR (f) resultado de aplicar AREVDCM.

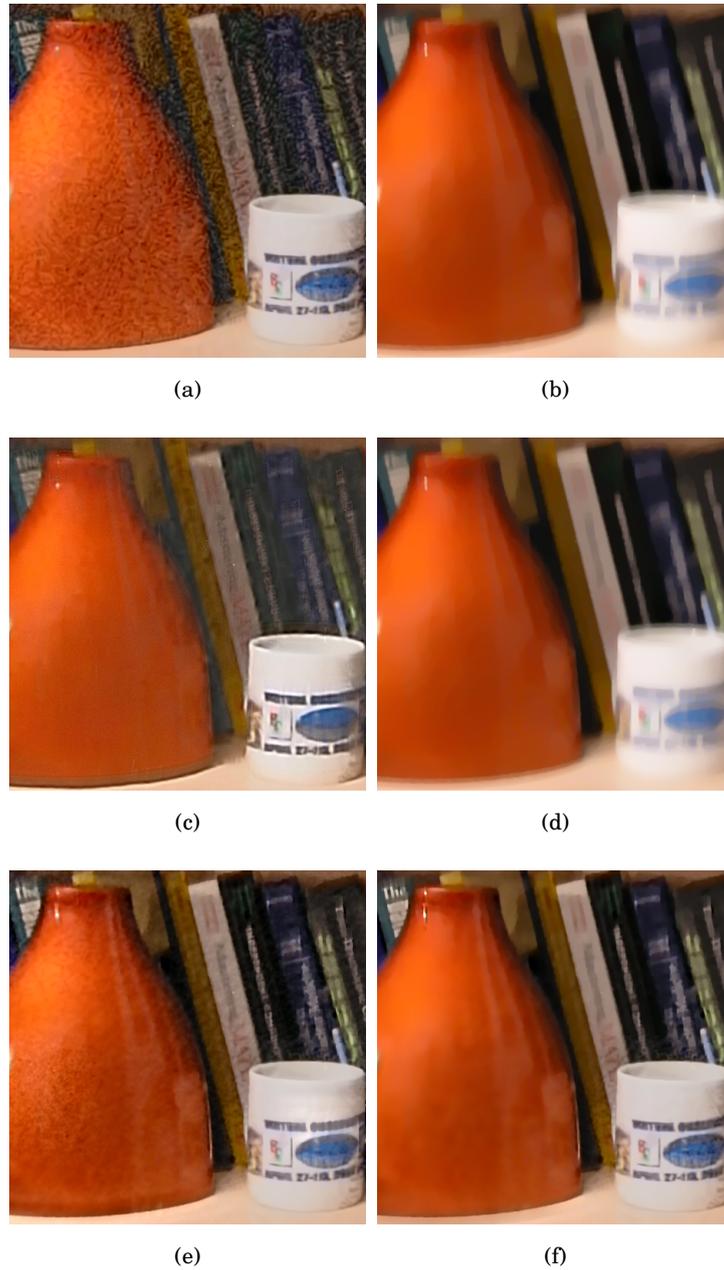


Figura 5.12: Resultados para el segundo experimento: (a) imagen tras supresión de ruido usando [1], (b) restauración con el método en [2] (c) imagen restaurada con el método [3], (d) resultado de aplicar AREI, (e) resultado de aplicar AREVDR (f) resultado de aplicar AREVDCM.

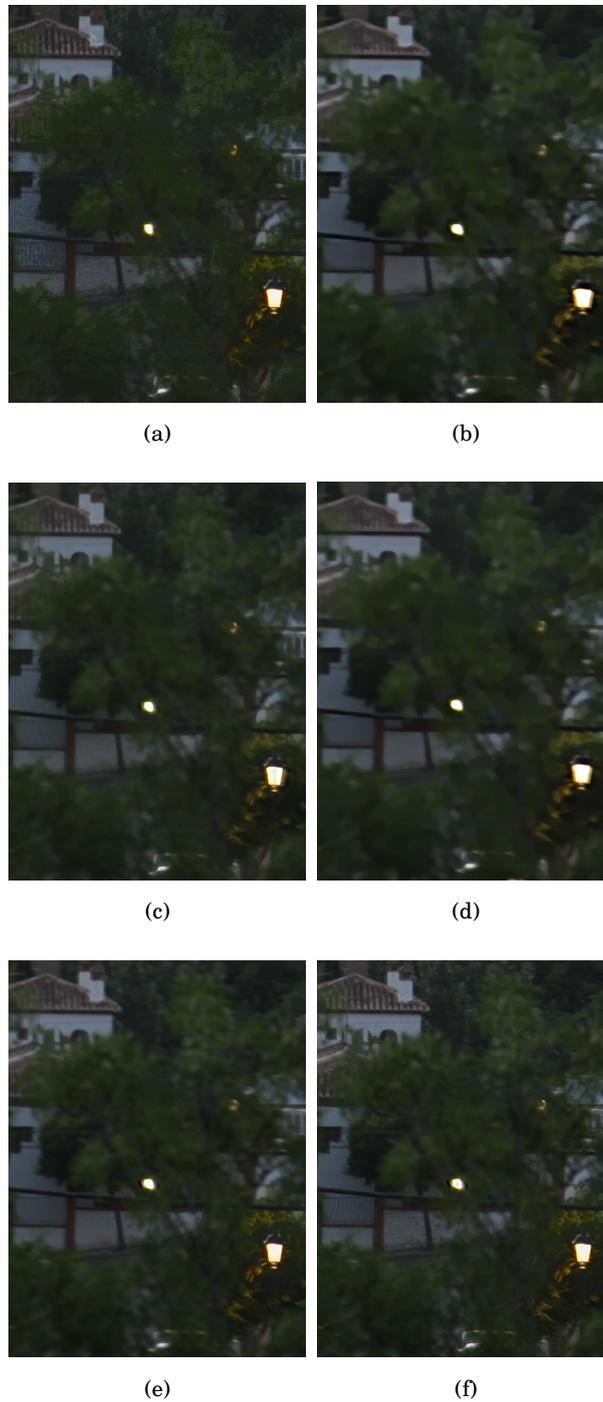


Figura 5.13: Resultados para el tercer experimento: (a) imagen tras supresión de ruido usando [1], (b) restauración con el método en [2] (c) imagen restaurada con el método [4], (d) resultado de aplicar AREI, (e) resultado de aplicar AREVDR (f) resultado de aplicar AREVDCM.

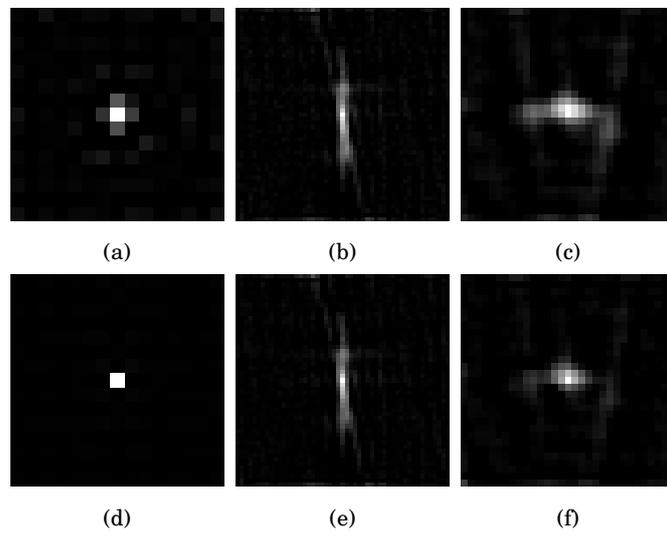


Figura 5.14: Núcleos de emborronamiento estimados por el método [2] (primera fila) y AREI del Capítulo 4 (segunda fila) para cada experimento real. Los valores de las PSFs fueron normalizados al rango $[0,255]$ por motivos de visualización. (a) y (d) Núcleo obtenido para las imágenes en las Figuras 5.10a y 5.10b, (b) y (e) PSF obtenida en el experimento de las Figuras 5.10c y 5.10d, (c) y (f) Núcleo de emborronamiento para las Figuras 5.10e y 5.10f.

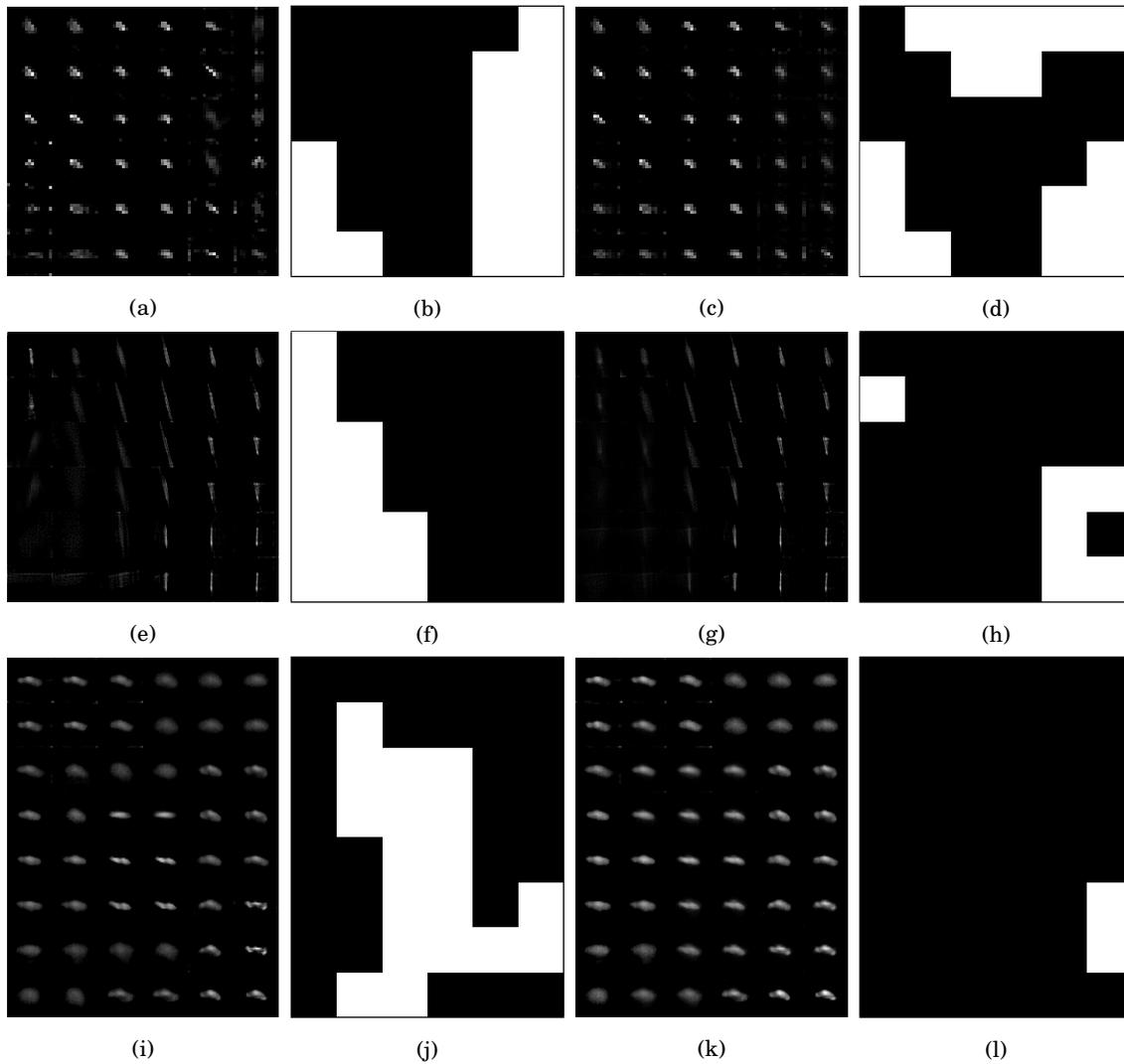


Figura 5.15: Núcleos de emborronamiento estimados y corregidos junto a la estrategia de selección para cada uno de los experimentos reales ejecutados sobre el AREVDR. Los bloques en blanco representan parches en los que se aplicó eliminación de ruido o corrección de PSF. Los valores de las PSFs fueron normalizados al rango $[0,255]$ por motivos de visualización. Cada fila corresponde con los experimentos en las Figuras 5.10a y 5.10b, 5.10c y 5.10d, 5.10e y 5.10f, respectivamente: (a),(e),(i) Núcleos obtenidos en cada experimento. (b),(f),(j) Bloques en los que se aplica la etapa de corrección. (c),(g),(k) PSFs corregidas. (d),(h),(l) Estrategia de selección donde se aplica el método en [1].

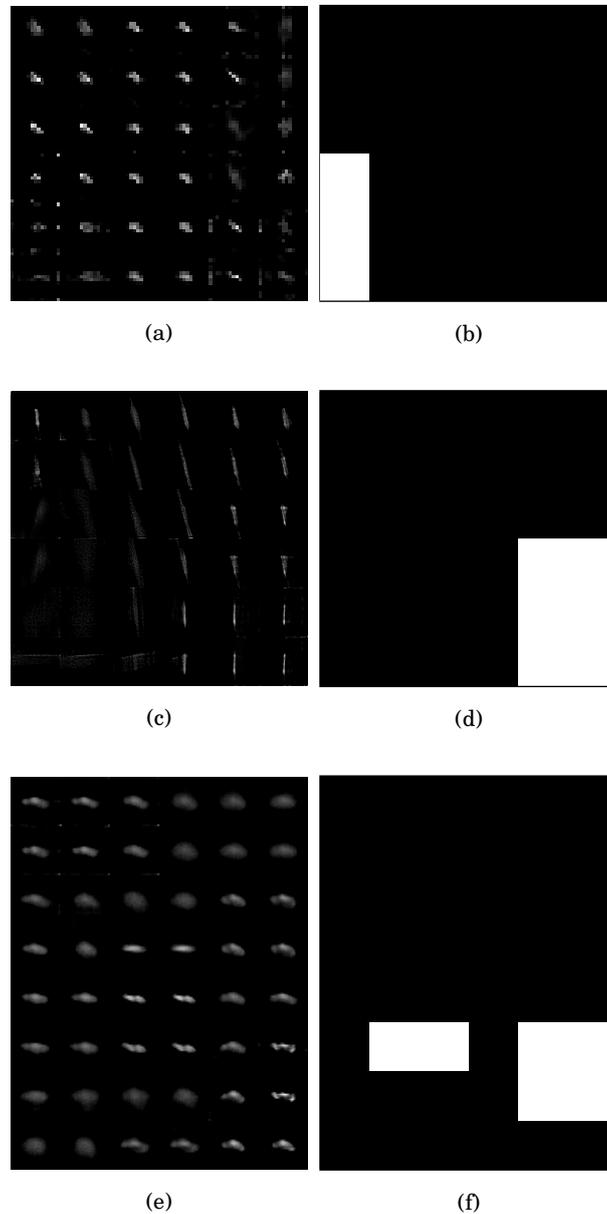


Figura 5.16: Núcleos de emborronamiento estimados y estrategia de selección para cada uno de los experimentos reales aplicando el AREVDCM. Los bloques en blanco representan parches en los que se aplicó eliminación de ruido. Los valores de las PSFs fueron normalizados al rango $[0,255]$ por motivos de visualización. (a) Núcleos obtenidos para las imágenes en las Figuras 5.10a y 5.10b. (b) Estrategia de selección para la restauración de la Figura 5.11f. (c) PSFs obtenidas en el experimento de las Figuras 5.10c y 5.10d. (d) Estrategia de selección para el experimento de la Figura 5.12f. (e) Núcleos de emborronamiento para las Figuras 5.10e y 5.10f. (f) Estrategia de selección en la restauración de la Figura 5.13f.

Capítulo 6

Conclusiones y trabajo futuro

*Una cosa es haber andado más camino y otra,
haber caminado más despacio.*

San Agustín (354-430).

Obispo y filósofo.

En esta memoria se planteó el problema de obtener imágenes nítidas en entornos de escasa luminosidad. El amplio conjunto de escenarios en los que puede ser útil despertó en nosotros gran interés. Sin embargo, presentaba una serie de dificultades que hubo que estudiar detenidamente y solucionar de manera aislada, para proceder con el proceso de restauración. Conforme avanzamos en la investigación, pudimos constatar que los métodos tradicionales que consideran un emborronamiento espacialmente invariante eran útiles en situaciones concretas pero limitados en cuanto a la restauración de escenas reales donde intervienen objetos en movimiento y a distintas profundidades. Para hacer frente a esta limitación consideramos un modelo de emborronamiento espacialmente variante más flexible y que proporcionó mejores resultados. A continuación exponemos las contribuciones más importantes de la memoria y las conclusiones que se extraen del trabajo realizado.

- Se ha hecho una revisión de los métodos de deconvolución ciega, tanto los que utilizan una única imagen como los que tratan el problema de exposición dual. A su vez se ha propuesto una taxonomía de los métodos de deconvolución a partir de pares de imágenes según el emborronamiento considerado. Destacamos varias razones por las que es provechoso la inclusión de la imagen de baja exposición al proceso de deconvolución. A pesar

de los problemas que ello conlleva, como necesidad de otra toma, calibración, etc., esta segunda imagen permite condicionar mejor el problema de deconvolución, facilitando tanto el modelado como la estimación más precisa del núcleo de emborronamiento y posibilita la obtención de una restauración más nítida.

- Se han analizado varias técnicas de calibración fotométrica y geométrica, y se ha propuesto un método robusto y automático de calibración de un par de imágenes de distinta exposición para la aplicación posterior del proceso de restauración.
- Se estudiaron los métodos de deconvolución ciega que combinan imágenes con distinta exposición y consideran un emborronamiento invariante y, en esta línea, se propuso un algoritmo de restauración que combina modelos de observación. El método propuesto presenta mejores resultados que los métodos actuales con que se comparó aunque, debido a la suposición de una estructura invariante del modelo de emborronamiento, no ofrece soluciones factibles para todos los casos de emborronamiento.
- Debido a las limitaciones de los métodos que consideran un modelo de emborronamiento invariante se proponen dos algoritmos de restauración que consideran un emborronamiento espacialmente variante. De los resultados obtenidos se deduce que el marco espacialmente variante es la forma óptima de resolver el problema considerado en esta memoria. Sin embargo, existe un compromiso entre el tamaño del emborronamiento y del bloque. Nótese que si el emborronamiento es severo habrá que considerar un tamaño de bloque mayor para poder llevar a cabo la estimación. Por ende, un menor número de bloques, podría afectar al potencial del marco propuesto.
- Los métodos propuestos mantienen una firme base teórica, además requieren de muy poca o ninguna supervisión del usuario, con una estimación automática de los parámetros del algoritmo resolviendo el problema de forma satisfactoria tanto en el dominio lineal del sensor como en la imagen JPEG revelada por la cámara. Sin embargo, los algoritmos desarrollados padecen de un coste computacional muy elevado para ser implementados en tiempo real.
- La aplicación de los métodos que se detallan en esta memoria no sólo beneficiaría a los usuarios de móviles, dónde los efectos del emborronamiento y ruido son aún más acusados por las dimensiones del sensor, sino que el marco propuesto se puede extender a otro

tipo de escenarios, tales como imágenes médicas, secuencias de videovigilancia, super-resolución del mapa de profundidad o creación de imágenes de alto rango dinámico HDR (*High Dynamic Range*).

Trabajos futuros

- Optimizar el Algoritmo 3 de restauración espacialmente variante por deconvolución rápida (AREVDR) orientado al desarrollo de una versión para móviles. En esta misma línea, realizar implementaciones masivamente paralelas de los algoritmos con el uso de GPUs.
- Estimación del parámetro λ (que controla la contribución de cada modelo de imagen) como otro parámetro desconocido del modelo jerárquico Bayesiano y ajuste dinámico del tamaño del bloque. Para la consecución de esto último podría utilizarse un método de segmentación, que aumentaría la flexibilidad a expensas de un aumento en el tiempo de ejecución.
- Elaboración de un método que permita la combinación de más de dos imágenes con varios propósitos. El primero de ellos, equivalente al objetivo tratado en la memoria, combinar las imágenes para la obtención de una imagen nítida. En segundo lugar, obtener una imagen de alto rango dinámico equivalente a la obtenida de la técnica HDR. En este último, nuestra solución evitaría propagar el emborronamiento de las imágenes con tiempo de exposición más largo. Además la mayoría de las cámaras actuales disponen de un mecanismo (*exposure bracketing*) que permite realizar varias tomas con distinta exposición en un único disparo, técnica muy útil para nuestro propósito.

Bibliografía

- [1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Trans. on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [2] S. D. Babacan, J. Wang, R. Molina, and A. K. Katsaggelos, “Bayesian blind deconvolution from differently exposed image pairs,” *IEEE Trans. on Image Processing*, vol. 19, no. 11, pp. 2874–2888, 2010.
- [3] M. Tico and K. Pulli, “Image enhancement method via blur and noisy image fusion,” in *IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 1521–1524.
- [4] M. Šorel, F. Šroubek, and J. Flusser, “Towards super-resolution in the presence of spatially varying blur,” in *Super-resolution imaging*, Milanfar Peyman, Ed., chapter 7, pp. 187–218. CRC Press, 2010.
- [5] M. Tallón, J. Mateos, S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Combining observation models in dual exposure problems using the Kullback-Leibler divergence,” in *European Signal Processing Conference (EUSIPCO)*, 2010, pp. 323–327.
- [6] M. Tallón, J. Mateos, R. Molina, and A.K. Katsaggelos, “Image prior combination in space-variant blur deconvolution for the dual exposure problem,” in *7th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2011, pp. 408–413.
- [7] M. Tallón, J. Mateos, S.D. Babacan, R. Molina, and A.K. Katsaggelos, “Space-variant kernel deconvolution for dual exposure problem,” in *19th European Signal Processing Conference (EUSIPCO)*, 2011, pp. 1678–1682.

-
- [8] M. Tallón, S.D. Babacan, J. Mateos, M.N. Do, R. Molina, and A.K. Katsaggelos, “Upsampling and denoising of depth maps via joint-segmentation,” in *EUSIPCO 2012 (20th European Signal Processing Conference 2012)*, 2012.
- [9] M.Tallón, J. Mateos, S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Space-variant blur deconvolution and denoising in the dual exposure problem,” *Information Fusion*, pp.—, 2012.
- [10] “Cambridge in colour,” <http://www.cambridgeincolour.com/>.
- [11] H. Rodríguez, *Captura digital y revelado RAW*, Marcombo, 2011.
- [12] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *SIGGRAPH ’97: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, 1997, pp. 369–378.
- [13] “Dcraw,” <http://www.cybercom.net/~dcoffin/dcraw/>.
- [14] A. K. Katsaggelos, *Digital Image Restoration*, Springer Series in Information Sciences, vol. 23, Springer-Verlag, 1991.
- [15] M. R. Banham and A. K. Katsaggelos, “Digital image restoration,” *IEEE Signal Processing Magazine*, vol. 14, no. 2, pp. 24–41, 1997.
- [16] T. G. Stockham, Jr., T. M. Cannon, and R. B. Ingebreetsen, “Blind deconvolution through digital signal processing,” *Proceedings IEEE*, vol. 63, no. 4, pp. 678–692, 1975.
- [17] D. Kundur and D. Hatzinakos, “Blind image deconvolution,” *IEEE Signal Processing Magazine*, vol. 13, no. 3, pp. 43–64, 1996.
- [18] D. Kundur and D. Hatzinakos, “Blind image deconvolution revisited,” *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 61–63, 1996.
- [19] T. E. Bishop, S. D. Babacan, B. Amizic, A. K. Katsaggelos, T. Chan, and R. Molina, “Blind image deconvolution: problem formulation and existing approaches,” in *Blind image deconvolution: theory and applications*, P. Campisi and K. Egiazarian, Eds., chapter 1. CRC press, 2007.

- [20] M. Vega, J. Mateos, R. Molina, and A. K. Katsaggelos, "Astronomical image restoration using variational methods and model combination," *Statistical Methodology*, vol. 9, no. 1-2, pp. 19–31, 2012.
- [21] I. Amro, J. Mateos, and M. Vega, "Parameter estimation in Bayesian super-resolution pansharpening using contourlets," in *IEEE International Conference on Image Processing ICIP 2011*. Brussels (Belgium), 2011, pp. 1345–1348.
- [22] M. Luessi, S.D. Babacan, R. Molina, J.R. Booth, and A.K. Katsaggelos, "Bayesian symmetrical EEG/fMRI fusion with spatially adaptive priors.," *NeuroImage*, vol. 55, no. 1, pp. 113–132, 2011.
- [23] S. Villena, M. Vega, D. Babacan, R. Molina, and A. K. Katsaggelos, "Using the Kullback-Leibler divergence to combine image priors in super-resolution image reconstruction," in *IEEE International Conference on Image Processing (ICIP)*. Hong-Kong (China), 2010, pp. 893–896.
- [24] S. Dai, M. Yang, Y. Wu, and A. K. Katsaggelos, "Tracking motion-blurred targets in video," in *International Conference on Image Processing (ICIP'06)*, Atlanta, GA, October 2006.
- [25] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind deconvolution algorithms," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 33, no. 12, pp. 2354–2367, 2011.
- [26] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph," *ACM Transactions on Graphics, SIGGRAPH 2006 Conference Proceedings, Boston, MA*, vol. 25, pp. 787–794, 2006.
- [27] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Transactions on Graphics, SIGGRAPH 2008 Conference Proceedings*, 2008.
- [28] S.D. Babacan, R. Molina, and A.K. Katsaggelos, "Variational Bayesian blind deconvolution using a total variation prior," *IEEE Trans. on Image Processing*, vol. 18, pp. 12 – 26, 2009.
- [29] H. W. Richardson, "Bayesian-based iterative method of image restoration," *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55–59, 1972.

-
- [30] L. B. Lucy, “An iterative technique for the rectification of observed distributions,” *Astronomical Journal*, vol. 79, pp. 745+, 1974.
- [31] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf, “Fast removal of non-uniform camera shake,” in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 463–470.
- [32] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, “Non-uniform deblurring for shaken images,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2010)*, 2010, pp. 491–498.
- [33] D. Krishnan and R. Fergus, “Fast image deconvolution using hyper-laplacian,” in *Advances in Neural Information Processing Systems (NIPS)*, 2009.
- [34] A. Buades, B. Coll, and J. M. Morel, “A review of image denoising algorithms, with a new one,” *Multiscale Model and Simulation*, vol. 4, pp. 490–530, 2005.
- [35] A. Buades, B. Coll, and J. M. Morel, “A non-local algorithm for image denoising,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, 2005, vol. 2, pp. 60–65.
- [36] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, “Image denoising using scale mixtures of Gaussians in the wavelet domain,” *IEEE Trans. on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.
- [37] L. Yuan, J. Sun, L. Quan, and H-Y. Shum, “Blurred/non-blurred image alignment using sparseness prior,” in *IEEE International Conference on Computer Vision (ICCV)*, 2007, pp. 1–8.
- [38] L. Yuan, J. Sun, L. Quan, and H-Y. Shum, “Image deblurring with blurred/noisy image pairs,” *ACM Trans. Graph.*, vol. 26, 2007.
- [39] E. Kee, S. Paris, S. Chen, and J. Wang, “Modeling and removing spatially-varying optical blur,” in *Proc. IEEE International Conference of Computational Photography (ICCP)*, 2011, pp. 1–8.

- [40] M. Šorel and F. Šroubek, "Space-variant deblurring using one blurred and one underexposed image," in *IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 157–160.
- [41] J. Jia, J. Sun, C-K. Tang, and H-Y. Shum, "Bayesian correction of image intensity with spatial consideration," *European Conference on Computer Vision (ECCV)*, pp. 342–354, 2004.
- [42] M. Tico, N. Gelfand, and K. Pulli, "Motion-blur-free exposure fusion," in *IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 3321–3324.
- [43] N. Cvejic, T. Seppanen, and S. J. Godsill, "A nonreference image fusion metric based on the regional importance measure," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 212–221, 2009.
- [44] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.
- [45] S. Babacan, R. Molina, and A. Katsaggelos, "Variational Bayesian super resolution," *IEEE Trans. on Image Processing*, vol. 20, no. 99, pp. 984–999, 2011.
- [46] P. Ruiz, J. Mateos, R. Molina, and A.K. Katsaggelos, "A Bayesian active learning framework for a two-class classification problem," in *MUSCLE International Workshop on Computational Intelligence for Multimedia Understanding*, E. Salerno, A. E. Cetin, and O. Salvetti, Eds. Pisa (Italy), 2012, vol. LNCS-7252, pp. 42–53.
- [47] G. Chantas, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Variational Bayesian image restoration with a product of spatially weighted total variation image priors," *IEEE Trans. on Image Processing*, vol. 19, no. 2, pp. 351–362, 2010.
- [48] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 2657–2664.
- [49] M. G. Kang and A. K. Katsaggelos, "General choice of the regularization functional in regularized image restoration," *IEEE Trans. on Image Processing*, vol. 4, no. 5, pp. 594–602, 1995.

- [50] R. Molina, A. K. Katsaggelos, and J. Mateos, “Bayesian and regularization methods for hyperparameter estimation in image restoration,” *IEEE Trans. on Image Processing*, vol. 8, no. 2, pp. 231–246, 1999.
- [51] J. Miskin, *Ensemble Learning for Independent Component Analysis*, Ph.D. thesis, Astrophysics Group, University of Cambridge, 2000.
- [52] M.J. Beal, *Variational algorithms for approximate Bayesian inference*, Ph.D. thesis, The Gatsby Computational Neuroscience Unit, University College London, 2003.
- [53] V. Šmídl and A. Quinn, *The Variational Bayes Method in Signal Processing*, Springer Verlag, 2006.
- [54] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, 2006.
- [55] M. I. Jordan, Z. Ghahramani, T. S. Jaakola, and L. K. Saul, “An introduction to variational methods for graphical models,” in *Learning in Graphical Models*, pp. 105–162. MIT Press, 1998.
- [56] S. Kullback and R. A. Leibler, “On information and sufficiency,” *Annals of Mathematical Statistics*, vol. 22, pp. 79–86, 1951.
- [57] S. Kullback, *Information Theory and Statistics*, New York, Dover Publications, 1959.
- [58] C.M. Bishop and M.E. Tipping, “Variational relevance vector machine,” in *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*. 2000, pp. 46–53, Morgan Kaufmann Publishers.
- [59] R. Molina, M. Vega, J. Mateos, and A.K. Katsaggelos, “Variational posterior distribution approximation in Bayesian super resolution reconstruction of multispectral images,” *Applied and Computational Harmonic Analysis, special issue on “Mathematical Imaging”, Part II*, vol. 24, no. 2, pp. 251–267, February 2008.
- [60] R. Molina, J. Mateos, and A. K. Katsaggelos, “Blind deconvolution using a variational approach to parameter, image, and blur estimation,” *IEEE Trans. on Image Processing*, vol. 15, no. 12, pp. 3715–3727, Dec. 2006.
- [61] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, second edition, 2004.

- [62] B. Zitová, J. Flusser, and F. Šroubek, “Image registration methods: a survey and recent advances,” in *IEEE International Conference on Image Processing (ICIP Tutorial)*, 2005.
- [63] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, 2008.
- [64] A. A. Goshtasby, *Image registration: Principles, tools and methods*, Springer, 2012.
- [65] J. Chen, L. Yuan, C.-K. Tang, and L. Quan, “Robust dual motion deblurring,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, pp. 1–8, June 2008.
- [66] D.-K. Lim and Y.-S. Ho, “A fast block-matching motion estimation algorithm with motion modeling and motion analysis,” in *Proceedings of the Third IEEE Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing*, London, UK, UK, 2002, PCM '02, pp. 135–142, Springer-Verlag.
- [67] H. Bay, T. Tuytelaars, and L. V. Gool, “Surf: Speeded up robust features,” in *European Conf. on Computer Vision (ECCV)*, 2006, pp. 404–417.
- [68] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [69] D. G. Lowe, “Object recognition from local scale-invariant features,” *IEEE International Conference on Computer Vision (ICCV)*, vol. 2, pp. 1150–1157, 1999.
- [70] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [71] G. Chantas, N. Galatsanos, R. Molina, and A. Katsaggelos, “Variational Bayesian inference image restoration using a product of total variation-like image priors,” in *Proc. 2nd Int Cognitive Information Processing (CIP) Workshop*, 2010, pp. 227–231.
- [72] F. Šroubek and J. Flusser, “Multichannel blind deconvolution of spatially misaligned images,” *IEEE Journal on Image Processing*, vol. 7, pp. 45–53, 2005.
- [73] S.-n. Cho and S. Lee, “Fast motion deblurring,” in *ACM Transaction on Graphics SIG-GRAPH*, New York, NY, USA, 2009, pp. 145:1–145:8, ACM.

- [74] S. D. Babacan, R. Molina, and A.K. Katsaggelos, “Parameter estimation in TV image restoration using variational distribution approximation,” *IEEE Trans. on Image Processing*, vol. 17, no. 3, pp. 326–339, 2008.
- [75] J. M. Bioucas-Dias, M. A. T. Figueiredo, and J.P. Oliveira, “Total variation-based image deconvolution: A majorization-minimization approach,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, 2006.
- [76] G. Parisi, *Statistical Field Theory*, Addison-Wesley, Redwood City, CA, 1988.
- [77] M. Tipping and C. Bishop, “Bayesian image super-resolution,” in *Advances in Neural Information Processing Systems 15*, S. Thrun, S. Becker, and K. Obermayer, Eds., Cambridge, MA, 2003, pp. 1279–1286, MIT Press.
- [78] A. Kanemura, S.-I. Maeda, and S. Ishii, “Superresolution with compound Markov random fields via the variational EM algorithm,” *Neural Networks*, vol. 22, no. 7, pp. 1025 – 1034, 2009.
- [79] “Microsoft Kinect for Windows,” <http://kinectforwindows.org/>.
- [80] A. Bovik, Ed., *Handbook of Image Video Processing*, Elsevier, second edition, 2005, Total Variation Regularization, pages 192, 193.
- [81] B. D. Ripley, *Spatial Statistics*, pp. 88–90, John Wiley, 1981, 1981.
- [82] S. Villena, M. Vega, R. Molina, and A. K. Katsaggelos, “Image prior combination in super-resolution image reconstruction,” in *European Signal Processing Conference (EUSIPCO 2010)*. Aalborg (Denmark), August 2010, pp. 616–620.
- [83] J.-L. Starck, M. Elad, and D. L. Donoho, “Image decomposition via the combination of sparse representations and a variational approach,” *IEEE Trans. on Image Processing*, vol. 14, no. 10, pp. 1570–1582, 2005.
- [84] M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling, “Efficient filter flow for space-variant multiframe blind deconvolution,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2010)*, 2010, pp. 607–614.