Universidad de Granada

Departamento de Ciencias de la Computación e Inteligencia Artificial

Programa de Doctorado en Tecnologías de la Información y la Comunicación

New Developments in Evolutionary Image Registration for Complex 3D Scenarios

Tesis Doctoral

Enrique Bermejo Nievas

Granada, Noviembre de 2017

UNIVERSIDAD DE GRANADA

New Developments in Evolutionary Image Registration for Complex 3D Scenarios

Memoria que presenta

Enrique Bermejo Nievas

Para optar al grado de Doctor

Noviembre de 2017

Directores

Oscar Cordón García
Manuel Chica Serrano

Departamento de Ciencias de la Computación e Inteligencia
Artificial

Agradecimientos

En primer lugar, mencionar que este trabajo ha sido posible gracias a la Universidad de Granada y a la financiación aportada por el Ministerio de Economía, Industria y Competitividad, bajo la concesión de la ayuda para contratos predoctorales para la formación de doctores (referencia BES-2013-062587), asociada al proyecto de investigación SOCOVIFI2 (referencia TIN2012-38525-C02-01).

Quiero agradecer profundamente a mis directores, Óscar y Manuel, por la oportunidad de trabajar en un equipo excepcional, todo un ejemplo de profesionalidad, rigor y cercanía personal. Ha sido un honor y un placer compartir este viaje con ellos y con el resto del grupo de investigación. Su disponibilidad, apoyo y amistad han sido incalculables y me han permitido llegar al final de esta etapa. Sin su confianza en mí ni su constante motivación, esta tesis no habría sido posible.

Aun así, no son las únicas personas que merecen agradecimiento, ya que esta tesis no habría sido posible sin el empujón inicial de José Santamaría o la ayuda de Inma y Fernando del Laboratorio de Antropología de la UGR. También a Sergio y a Sancho por la calurosa acogida en sus respectivos grupos de trabajo durante mis estancias en el European Centre for Soft Computing de Asturias y la Universidad de Alcalá. Agradezco enormemente la oportunidad de aprender de profesionales tan agradables.

A mis padres y hermana por su cariño y confianza incondicional. A toda la gente de Ciudad Real, con la que di los primeros pasos en el mundo de la investigación, y de Granada, con la que he compartido momentos inolvidables. Por último, a Ana. *Chiar nu ştiu cum să îţi mulţumesc pentru răbdarea ta, sprijin şi afecţiune. Mulţumesc, jumatatea mea.* En esta tesis está reflejado todo vuestro apoyo y amistad, a pesar de la distancia.

El doctorando / The *doctoral candidate* **[Enrique Bermejo Nievas]** y los directores de la tesis / and the thesis supervisor/s: **[Oscar Cordón García y Manuel Chica Serrano]**

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.
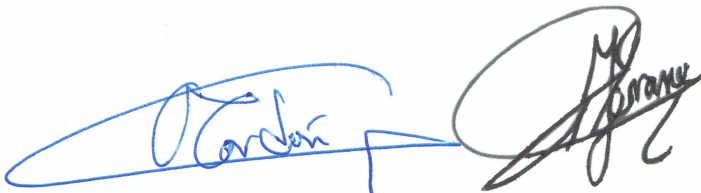
/

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / Place and date:

Granada, 21 de Noviembre de 2017

Director/es de la Tesis / *Thesis supervisor/s;*   Doctorando / *Doctoral candidate*:

Firma / Signed   Firma / Signed

Contents

Resumen

1 Introducción al problema

El desarrollo de las tecnologías digitales en 3D ha desembocado en notables avances en diferentes áreas como la robótica, medicina, astrofotografía o las industrias automovilística y del entretenimiento. Impulsados por estos avances, diferentes áreas de investigación han cobrado una gran relevancia dada la necesidad de representar el mundo que nos rodea en tres dimensiones y analizar su contenido. En concreto, el *registrado de imágenes* [Gos05] es una tarea indispensable en *visión por computador*, a medio camino entre las técnicas de procesamiento y de análisis de imágenes, con un claro propósito: extraer información relevante de imágenes digitales.

El registrado de imágenes consiste en la búsqueda de una transformación óptima entre dos o más imágenes provenientes de orígenes diferentes. El principal objetivo reside en establecer un sistema de coordenadas común que dé lugar a una correspondencia espacial entre las imágenes en cuestión. Para ello, la búsqueda de la transformación óptima se formula como un problema de optimización. El registrado permite relacionar la información contenida en conjuntos de imágenes desde diferentes ángulos, por ejemplo, integrando información obtenida usando diferentes sensores o dispositivos de adquisición (análisis multimodal), reconstruyendo información tridimensional desde múltiples perspectivas u objetos móviles (análisis multivista), identificando imágenes usando enfoques basados en modelos (correspondencia de plantillas) o evaluando cambios en escenas capturadas en diversos momentos (análisis multitemporal).

Junto con el desarrollo de las tecnologías 3D, las técnicas de registrado de imágenes también han evolucionado haciendo posible encontrar métodos versátiles, eficientes, robustos y capaces de operar en tiempo real. Sin embargo, estos criterios únicamente se cumplen en condiciones ideales o escenarios controlados. De hecho, la imprecisión asociada a cada contexto particular y los exigentes requisitos computacionales de la gran mayoría de aplicaciones reales hacen imposible el diseño de métodos de registrado universales.

La eficacia y el rendimiento de las técnicas de registrado se ven afectados por una gran variedad de factores. En el marco de esta tesis doctoral, se ha hecho hincapié en la identificación de diferentes factores decisivos que influyen en la complejidad de dichas técnicas. En concreto se ha abordado el problema de registrado en tres aplicaciones diferentes. En primer lugar, la reconstrucción de objetos en 3D considerando imágenes de rango adquiridas mediante una cámara de tiempo de vuelo. Debido a la baja calidad y resolución de las imágenes obtenidas mediante este tipo de dispositivos, se plantea el desarrollo de un método para la mejora de la calidad de imágenes de rango junto con un novedoso método de registrado de imágenes basado en *inteligencia de enjambres*. A continuación, este método bio-inspirado se extiende a un problema de imagen médica en 3D consistente en la alineación de resonancias magnéticas cerebrales. Para finalizar, se pretende abordar un problema de identificación forense. En concreto, se plantea el modelado del movimiento mandibular y su integración en el diseño de un método de superposición craneofacial.

Este problema implica una gran complejidad al requerir el registrado 3D/2D de un modelo craneal articulado superpuesto sobre la fotografía ante mortem de una persona desaparecida.

Para lograr los objetivos planteados en esta contribución, se han empleado herramientas de *soft computing* [Zad94, TT08] e *inteligencia artificial*. Dada la complejidad de los problemas abordados, se han considerado técnicas heurísticas englobadas en campos como la computación evolutiva [ES03] y la inteligencia de enjambres [KES01]. Estas técnicas han adquirido una gran relevancia por su habilidad para manejar imprecisión y ambigüedad a la vez que proporcionan resultados excelentes de forma eficiente.

2 Desarrollo realizado

El desarrollo de la tesis puede dividirse en tres bloques diferenciados. En cada bloque se realiza un análisis de la complejidad asociada a cada una de las aplicaciones abordadas en esta contribución. Además, se detallan las propuestas diseñadas para abordar los factores desencadenantes de dicha complejidad y solventar las dificultades que presentan.

2.1 Registrado de imágenes de rango

La representación tridimensional de objetos cotidianos es un tema candente en visión por computador. Diferentes tareas como la reconstrucción 3D, el modelado y el reconocimiento de objetos juegan un papel esencial en una gran variedad de aplicaciones. La adquisición de imágenes y su alineación son componentes indispensables de dichas tareas. Generalmente, el proceso se inicia muestreando un escenario cualquiera y capturando imágenes 3D mediante un escáner láser o de luz estructurada, un sistema de visión estereoscópica, una cámara de tiempo de vuelo o usando técnicas como la fotogrametría a partir de varias imágenes 2D. El primer factor limitante suele aparecer en esta etapa, debido a que la gran mayoría de dispositivos posee una serie de características y restricciones como resolución, campo de visión, precisión, etc.

Para esta aplicación se ha considerado un tipo concreto de dispositivo 3D, conocido como *cámara de tiempo de vuelo*. Su funcionamiento se basa en la medición del tiempo que tarda un haz de luz en viajar desde un sensor emisor de la cámara hasta un sensor receptor, reflejándose en la superficie de los objetos de una escena. Este tipo de dispositivos se caracteriza por alcanzar una tasa de fotogramas muy elevada, permitiendo grabar video en tiempo real, pero se encuentran limitados por diversos problemas que derivan en imágenes de baja calidad y resolución, con altos niveles de ruido e imprecisión.

Debido a las características de las imágenes de tiempo de vuelo, hemos desarrollado un método para mejorar su calidad que se divide en tres fases que pretenden reducir la cantidad de ruido y facilitar la extracción de información relevante. En primer lugar, se realiza una calibración para reducir el efecto de las condiciones de iluminación en la medición de las distancias de profundidad de la cámara, acompañada del promedio de un número reducido de fotogramas con el fin de minimizar los errores ocasionados por inter-reflexión del haz de luz. De esta forma se atenúa la deformación en la superficie de los objetos debido a errores de captura. Tras ello, se aplica un método para corregir la iluminación de la imagen y una técnica de mejora basada en forma a partir de sombra, o *shape from shading* [ZTC99]. Con esta etapa se corrige una gran cantidad de ruido y se suaviza la superficie de los objetos capturados. Finalmente, se extrae un conjunto de características que permitirá realizar el registrado de dos vistas adyacentes de un objeto para su reconstrucción. En concreto, se aplican dos descriptores de características, analizando la curvatura [CCFM08] y la distribución de intensidades [BETV08] que definen la superficie de los objetos.

Una vez se ha mejorado la calidad de las imágenes y extraído la información más relevante, se procede a realizar la alineación de las imágenes. Para ello nos proponemos emplear un novedoso

método bio-inspirado basado en la alimentación de bacterias, conocido como *Bacterial Foraging Optimization Algorithm* (BFOA). Este algoritmo se basa en el comportamiento de una bacteria presente en nuestro organismo, Escherichia Coli, durante la búsqueda de nutrientes. En particular, hemos diseñado 52 variantes diferentes sobre la estructura del algoritmo original [Pas02] y realizado una experimentación exhaustiva considerando diferentes escenarios complejos generados mediante una cámara de tiempo de vuelo.

2.2 Registrado de imágenes médicas

En este apartado se aborda una aplicación complementaria, el registrado de imágenes médicas [HHH01, PMV03]. En concreto, un problema caracterizado por la gran cantidad de información incluida en imágenes obtenidas mediante técnicas de resonancia magnética. Generalmente, los métodos de registrado (conocidos como métodos basados en intensidades) emplean métricas de correlación para comparar los patrones de intensidades de las imágenes. En este caso la complejidad del problema reside en que la gran cantidad de información asociada a este tipo de imágenes médicas excede las capacidades computacionales disponibles y el tiempo requerido por los métodos de registrado impide su aplicabilidad en situaciones reales.

Así, se propone aplicar una reducción de la dimensionalidad del problema empleando nuevos métodos de extracción de características. Debido a las particularidades de esta aplicación, dichos métodos son propensos a la pérdida de información relevante, complicando aún más el problema de registrado. Por ello, hemos extendido el método de registrado bio-inspirado diseñado anteriormente. En esta ocasión, consideramos un enfoque basado en características como una etapa de prealineación destinada a encontrar una aproximación a la transformación de registrado óptima de forma eficiente.

En particular, se han evaluado las diferentes variantes propuestas para el método bio-inspirado en un problema de registrado considerando imágenes cerebrales con diferentes niveles de ruido y lesiones, simulando un enfoque multi-temporal. Generalmente, la superficie cortical o surcos cerebrales se analizan con el fin de extraer información relevante de imágenes cerebrales. En nuestra propuesta, hemos considerado un detector de líneas de cresta [YBS05] que nos permitirá extraer información anatómica a través de la curvatura de los pliegues cerebrales manteniendo una alta eficiencia en el proceso.

2.3 Superposición craneofacial

Existen numerosas técnicas forenses que permiten conocer la identidad de un individuo, como el reconocimiento de huellas dactilares, las muestras de ADN, o los historiales dentales [Bur99, YS00]. En algunas situaciones como pueden ser desastres naturales, crímenes de guerra, o desapariciones, muchas de estas técnicas no son aplicables. La *superposición craneofacial* [AISB95, Yos12] es una técnica alternativa que permite realizar el análisis forense en aquellos casos donde únicamente se recuperan restos óseos. Dado que su aplicación es tediosa y está basada en un proceso subjetivo que requiere horas de procesamiento manual, recientemente se han propuesto diferentes métodos con el propósito de automatizar esta tarea [IBC+09, ICD12, CÁInCA+15].

La técnica de superposición craneofacial considerada implica el registrado de un modelo craneal en 3D con una fotografía ante mortem en 2D. La relación entre ambas imágenes se obtiene mediante la combinación de una proyección con una transformación rígida, que permiten determinar la pose del modelo 3D relativa a la localización de la cámara en la fotografía. Una vez alineadas, un antropólogo forense se encarga de estudiar las características morfológicas de ambas imágenes analizando la similitud entre dos conjuntos de puntos: landmarks cefalométricos en la fotografía facial y landmarks craneométricos en el modelo craneal. Puesto que este problema implica la correspondencia de dos objetos diferentes, el proceso de registrado se convierte en una tarea

de optimización compleja guiada por información incompleta e imprecisa, generando un extenso espacio de búsqueda con múltiples mínimos locales.

Los métodos automáticos de superposición propuestos hasta la fecha parten de un modelo craneal rígido en 3D, obtenido generalmente escaneando el cráneo junto a la mandíbula en una postura aproximada a la pose facial en la fotografía ante-mortem. Desde el punto de vista anatómico, esta simplificación afecta a la precisión y fiabilidad del método de registrado, además de inducir identificaciones incorrectas. Por otro lado, uno de los criterios empleados por expertos forenses para evaluar la calidad de la superposición es la correspondencia entre las piezas dentales del cráneo y de la fotografía cuando son visibles.

Con el fin de extender y mejorar el actual estado del arte, se plantea considerar el modelo craneal como un objeto articulado. Así, hemos analizado y modelado la articulación mandibular en condiciones normales, teniendo en cuenta las restricciones anatómicas involucradas en su movimiento. Además, hemos rediseñado el método actual de superposición craneofacial integrando el modelo 3D de la mandíbula como un objeto móvil, independiente del cráneo.

Debido a la novedad de este enfoque, hemos diseñado un estudio experimental generando un conjunto de escenarios de superposición craneofacial donde la mayoría de fuentes de imprecisión durante la apertura mandibular están controladas. Este banco de pruebas constituye un método objetivo de evaluación para nuestra propuesta de superposición articulada, permitiendo analizar los efectos del modelo mandibular a diferentes aperturas.

3   Conclusiones y Trabajos Futuros

La relevancia de los métodos de registrado en aplicaciones cotidianas es indiscutible, como se refleja en la frenética actividad relacionada con el registrado de imágenes en áreas como la visión o los gráficos por computador. La complejidad inherente en los problemas de registrado hace que sea considerado como un campo de investigación dinámico, cuyo objetivo es encontrar métodos fiables y eficientes que a la vez produzcan soluciones precisas y robustas. La principal motivación de esta tesis doctoral reside en la identificación de diferentes factores limitantes en aplicaciones reales, así como en el diseño de técnicas novedosas que permitan superar estas dificultades. Recientemente, diferentes técnicas de soft computing e inteligencia artificial, como metaheurísticas, algoritmos meméticos o métodos bio-inspirados, han demostrado su eficacia abordando problemas caracterizados por su gran complejidad.

En esta tesis doctoral se presentan dos contribuciones principales, adoptando los últimos avances en campos como soft computing y visión por computador. La primera propuesta está relacionada con el empleo de un método de inteligencia de enjambres, basado en el proceso de búsqueda de nutrientes de las bacterias, para abordar la complejidad existente en dos problemas de registrado de imágenes diferentes. En primer lugar, hemos realizado un estudio para evaluar la viabilidad de las cámaras de tiempo de vuelo en un problema de reconstrucción de objetos en 3D. Debido a la complejidad inducida por la baja calidad de las imágenes adquiridas mediante este tipo de dispositivos, se ha desarrollado un método de mejora de calidad para imágenes de rango. Dicho método minimiza el ruido producido durante la captura de imágenes e integra un descriptor de características que permite extraer información relevante de la superficie de los objetos combinando la información 2D junto con la geometría 3D de la escena. En cuanto al método de registrado bio-inspirado, hemos introducido una serie de modificaciones en el diseño original del algoritmo basado en bacterias. Con el fin de mejorar el estado del arte, se han diseñado 52 variantes diferentes a partir de la combinación de dichas modificaciones. Las diferentes variantes se han adaptado al problema de reconstrucción 3D siguiendo un enfoque por pares que permite alinear vistas adyacentes de los objetos.

Además, se ha planteado analizar las posibilidades de nuestra propuesta bio-inspirada en una aplicación complementaria, el registrado de imágenes médicas en 3D. En concreto, hemos considerado un problema exigente que implica la alineación de imágenes cerebrales adquiridas mediante técnicas de resonancia magnética. La complejidad de este problema está localizada en la gran cantidad de información existente en dichas imágenes, que además contienen lesiones y diferentes niveles de ruido. Para abordar este problema, hemos seguido un enfoque basado en características que permite reducir la dimensionalidad del problema, explotando la información de curvatura contenida en los pliegues de la superficie cerebral.

Se han llevado a cabo dos estudios experimentales para evaluar la idoneidad de las diferentes variantes propuestas, analizando el problema de reconstrucción 3D y de registrado de imágenes médicas. A partir de los resultados obtenidos, consideramos que un gran número de las variantes propuestas son capaces de mejorar considerablemente el diseño original del algoritmo de alimentación de bacterias. Aunque en algunos escenarios médicos nuestras variantes no han sido capaces de mejorar nítidamente los resultados del estado del arte actual, las alternativas de diseño y la alta modularidad del algoritmo considerado justifican su consideración como un método de registrado estable y aplicable a aplicaciones reales de gran complejidad.

La segunda contribución de esta tesis doctoral está relacionada con la complejidad existente en un problema concreto de registrado 3D/2D, la superposición craneofacial. En anteriores publicaciones [IBC$^+$09, ICDS11, CÁInCA$^+$15] se identificaron una serie de fuentes de imprecisión que afectan a la precisión de los métodos de superposición. En nuestra propuesta, hemos abordado una de estas fuentes con el objetivo de mejorar la robustez y fiabilidad de esta técnica de identificación. En concreto, hemos modelado e integrado la articulación mandibular en el diseño del método de superposición más reciente, propuesto en [CÁInCA$^+$15]. La idoneidad de este nuevo método se ha evaluado mediante un diseño experimental minucioso. Para ello, hemos diseñado una base de datos de casos de superposición que nos permite realizar la evaluación del modelo de apertura mandibular de forma objetiva. Los resultados obtenidos han puesto de manifiesto que la integración del modelo de apertura mejora considerablemente la fiabilidad, la precisión y la versatilidad del actual método del estado del arte en superposición craneofacial. Esta mejora considerable facilita el uso de fotografías en las que el individuo aparece sonriendo o con diferentes expresiones faciales donde la boca aparece parcialmente abierta. Este tipo de fotografías son generalmente descartadas en escenarios de identificación reales, por lo que la disponibilidad de nuevas evidencias ante mortem es una consecuencia directa del método desarrollado en esta contribución. La fiabilidad de la técnica de identificación forense basada en superposición craneofacial está directamente relacionada con la cantidad de fotos analizadas para un mismo individuo [DWK$^+$15, YIMS95, ASM94].

Recientemente han surgido una gran variedad de algoritmos bio-inspirados que destacan por su buen comportamiento tanto explorando como explotando eficientemente el espacio de búsqueda en problemas de alta complejidad. Como trabajos futuros, creemos que el estudio de estos nuevos enfoques en diferentes aplicaciones como el registrado de imágenes es beneficioso y conveniente para analizar el alcance de este tipo de métodos. Además, la tecnología de adquisición de imágenes en 3D ha evolucionado enormemente en los últimos años, dando lugar a una gran variedad de dispositivos caracterizados por su precisión, velocidad y portabilidad. Este desarrollo también ha influenciado a las técnicas involucradas en el proceso de reconstrucción 3D. Resultaría de gran interés realizar un estudio comparativo de las propiedades de los nuevos dispositivos disponibles, especialmente relevantes en aplicaciones como medicina o antropología forense donde se exigen modelos de alta calidad.

Con respecto al método articulado de superposición craneofacial, los resultados del modelo de apertura mandibular muestran un ligero error en la predicción de grandes aperturas. No obstante, consideramos que el modelo de regresión empleado satisface los requisitos del problema de

superposición, ya que las imágenes con grandes aperturas no se consideran en casos de identificación reales debido a las deformaciones del tejido facial. Además, los resultados de la integración del modelo mandibular en el método de superposición han sido notablemente satisfactorios tanto desde el punto de vista del proceso de optimización como de la calidad del solapamiento. Sin embargo, uno de los posibles trabajos futuros consistiría en analizar el impacto en la fiabilidad de la superposición de otro tipo de modelos articulares más complejos o realistas. Así mismo, planteamos el desarrollo de otro tipo de diseños innovadores que mejoren el proceso de optimización de la técnica de superposición.

A modo de resumen, concluimos que los resultados obtenidos en las tres aplicaciones diferentes corroboran el cumplimiento de los objetivos propuestos para esta tesis doctoral. Los resultados de la investigación desarrollada en este trabajo se han reflejado en tres publicaciones científicas en revistas internacionales indexadas en el JCR y un congreso internacional. Durante el desarrollo de esta disertación, se ha puesto de manifiesto la necesidad de considerar un enfoque polifacético a la hora de abordar problemas reales de aplicaciones interdisciplinares. No sólo es indispensable el desarrollo y la aplicación de métodos computacionales innovadores que permitan afrontar la complejidad de estos problemas, sino también considerar el amplio abanico de instrumentos y técnicas que proporcionan campos de investigación muy diversos.

Abstract

It can be argued that *image registration* is a ubiquitous *computer vision* task, playing a crucial role in a wide variety of problems from industrial to clinical applications. In essence, the registration task involves the search for a correspondence between images acquired under different circumstances. Relating the information of different sets of imaging data is essential for the application of other subsequent image analysis or computer vision tasks involving its integration, comparison, or manipulation. Along with the latest developments of three-dimensional imaging technology, image registration techniques have considerably evolved to a point where efficiency, versatility, robustness, and real-time performance are feasible goals. Nevertheless, these criteria are only fulfilled in ideal conditions or controlled scenarios. In fact, the intensive computational requirements and the unavoidable imprecision present in real world applications prevent a successful accomplishment of the previous goals.

Within the framework of this PhD thesis, the focus has been placed on identifying determinant factors leading to complexity in image registration domains that can be handled using *soft computing* and *artificial intelligence*. Both fields provide a huge number of techniques that have thrived in solving complex optimization problems. The relevance of such techniques relies on their ability to endure imprecision while reaching reasonable solutions efficiently in evolving environments where noise, incomplete or ambiguous information is present. The main motivation is thus to exploit the latter advantages to address a multidisciplinary IR problem handling the implicit complexity of recent and significant applications. Recently, different metaheuristic approaches, such as *evolutionary computation* and *swarm intelligence* methods belonging to a family of nature-inspired algorithms, have awakened a growing interest of the research community due to their promising capabilities to efficiently handle complex environments.

Several adverse conditions affect the performance and effectiveness of image registration techniques. In this contribution, different factors have been identified and assessed in order to overcome the difficulties of three distinct applications. In particular, we have addressed a 3D reconstruction problem involving time-of-flight range images. As a result of the low quality and resolution provided by these novel devices, we designed an image preprocessing pipeline to improve their quality along with a novel swarm intelligence registration method to face this challenging problem. Moreover, the proposed nature-inspired method has been extended to face a 3D medical image registration problem involving the alignment of 3D brain magnetic resonance scans. Finally, soft computing techniques have also been applied to address a forensic identification problem. Precisely, we designed a craniofacial superimposition method involving the articulation of the mandible. This problem is solved by applying a skull-face overlay technique, based on the 3D/2D registration of the articulated skull and the photograph of a missing individual.

Part I. Report

1 Statement

This section is devoted to describe the characteristics of the problems addressed in the current doctoral dissertation. After an introduction of the main aspects of image registration, we provide a justification for this work. Then, we enumerate the objectives scheduled for the development of the PhD dissertation. Last, the structure of the dissertation is shown.

1.1 Introduction

The major developments of imaging technology in the last two decades have led to remarkable breakthroughs in many areas such as robotics, medicine, automotive engineering, astrophotography, or entertainment industry. Pushed by these technological advancements, *image analysis* and *computer vision* (CV) [JH00], among others, have become especially relevant research fields while the challenges shifted from image acquisition to gathering and analyzing content information. Particularly, *image registration* (IR) [Gos05] or image alignment is an essential task that lies at the intersection of image processing and image analysis. The frontier between these fields is not clearly defined, as both share an ultimate goal: extracting relevant knowledge from digital images. IR allows the comparison and integration of imaging data from different sources (e.g. multiple photographs, different sensors, times, or viewpoints) [WEF+08]. This task has been successfully applied to several problems involving different sets of data, such as noise reduction [BCM05], image classification [TM09], image segmentation [PXP00], and 3D representation of real-world scenarios [BR02]. Even though IR itself is not the end-goal but instead the first stage of many applications, its relevance and contribution to *computer graphics* (image manipulation), *computer vision* (image analysis) and *artificial intelligence* (machine perception) is undeniable.

    IR is a process that aims at finding an optimal transformation between two or more images acquired under different conditions. The goal is to establish a common coordinate system that brings spatial correspondence between the considered images. In order to achieve this goal, the search for the optimal correspondence is usually treated as an optimization problem. Overall, IR is able to tackle different problem areas: information integration when images are provided by different acquisition devices or sensors (multimodal analysis), reconstruction of three-dimensional information from different viewpoints or moving objects (multiview analysis), image identification using model-based object recognition approaches (template matching), and change evaluation in scenes where images are acquired at different times (multitemporal analysis) [CF01, LRG+00, DCS11].

    A significant number of noteworthy IR algorithms have been presented so far. Classical methods are usually the most efficient and effective registration methods when the conditions are favorable and the problem scenarios are straightforward. However, ordinary registration algorithms

fail to produce accurate results when dealing with real-world applications where problem scenarios pose important challenges. IR is an extremely complex problem due to the large degree of variability of the imaging data. Classical numerical optimization techniques, although widely used, are known to present many limitations. Specially, the performance of these techniques is affected when images present noise, discretization, or a large amount of misalignment. In essence, classical methods provide a locally optimal solution and are prone to deliver inaccurate registrations.

Ideally, any IR method should reach real-time performance, while being computationally efficient, highly robust, and versatile. In reality, the technological advancements in image acquisition devices have created a dynamically evolving environment which hinders the quest of this goal. To overcome these difficulties, the use of *metaheuristics* (MHs) [Glo03] has been considered as global optimization techniques, although they are not free of the local minimum issue. These methods are alternative optimization approaches combining basic heuristic methods in higher-level procedures. Furthermore, *soft computing* (SC) techniques have a special relevance for this sort of complex problems as they are able to deal with imprecision and uncertainty in an efficient and reliable way [Zad94, TT08]. Particularly, biologically inspired MHs and their hybridizations, which are part of the *evolutionary computation* (EC)[BFM97, ES03] family, have proven to be efficient and accurate methods to explore the space of possible solutions, even when the optimization function is highly non-convex which is the case of many IR scenarios [SCD11, DCS11].

This dissertation is focused in a particular sort of IR problems: complex scenarios involving three-dimensional imaging data. The main motivation lies in the relevance of the digitization of 3D real world objects, which is an extremely fast-growing field. In fact, the ultimate goal of this work is to address a multidisciplinary problem in which novel and robust *artificial intelligence* (AI) techniques are applied to tackle the IR problem in complex scenarios within real-world applications. Particularly, three different areas of application have been assessed considering the important role of IR.

Not long time ago, 3D capturing sensors were inaccurate, slow, expensive, or bulky artifacts, while CV techniques for 3D reconstruction (from bi-dimensional images) required significant error-prone manual interaction (CAD-based approaches) and many processing steps [DF01], resulting in slow and inefficient approaches. Recently, diverse technologies have arisen and currently a wide variety of modern devices are available: from handheld 3D scanners to smartphones with integrated 3D cameras, able to provide high quality 3D images in real time. Moreover, reconstruction techniques have improved considerably and can operate automatically and efficiently thanks to the available computing power [CF01]. Range images are usually represented in two basic forms: a point cloud containing a list of 3D coordinates representing the scene, or a depth map relating the surface distances of every point in the scene to the camera viewpoint. 3D acquisition technology plays a crucial role in many relevant applications, particularly in industrial design, healthcare, cultural heritage, cinema, topography or architecture.

Modern medicine relies in the application of imaging tests for medical diagnosis and treatment. The revolutionary capabilities of imaging technology, along with recent developments in anatomical reconstruction, visualization, and analysis of 3D medical volumetric data have provided new opportunities for clinical applications [Sue09]. Many 3D image techniques or modalities are available [EJ16]: computed topography (CT), cone-beam CT (CBCT), magnetic resonance imaging (MRI), positron emission tomography (PET), and 3D ultrasound are the most renowned. 3D image processing (IP) have provided an extensive set of tools benefiting different clinical disciplines, with applications in diagnosis and disease monitoring [HRG$^+$09], intervention and treatment planning [SvdHK$^+$09], assisted surgery [MFW$^+$97, KRM$^+$10], anatomical atlas comparison [FF98, GDMW04], and anatomy segmentation [CE97, FG13, SAS13], among many others.

Forensic anthropology is a sub-field of physical anthropology that involves applying scientific knowledge to the collection and analysis of medico-legal information. As a discipline, it includes the recovery, identification, and description of human skeletal remains [Bur99, YS00]. Currently, several forensic identification techniques are available, i.e., DNA samples, fingerprint recognition, or dental records. When none of the previous methods can be applied, the analysis of skeletal remains becomes the last resort of forensic identification. One of the alternative techniques is craniofacial superimposition (CFS) [AISB95, Yos12], which involves the superimposition of a complete skull (or a skull model) with a number of *ante mortem* images of an individual. CFS is thus a technique used when only skeletal information is available for the forensic assessment and other techniques such as DNA or dental record analysis are not possible or conclusive [Isc81]. Since its definition on the nineteenth century, CFS has demonstrated to be a robust alternative for identification, solving an extensive number of cases: war crimes, mass disasters, missing people, etc. Nevertheless, its application is tedious and based on a 'trial and error' process requiring several hours of manual processing to obtain a correct superimposition. Therefore, there is a strong interest on designing automatic methods to assist the CFS identification procedure [IBC+09, DCI+11, ICDS11, ICD12].

1.2 Justification

Despite all the recent progress regarding imaging technology and CV algorithms, there are still many challenges ahead in the IP road map. Every improvement achieved in terms of hardware innovation, noise reduction, or quality enhancement is accompanied by an increased amount of data to be processed. Therefore, IR methods have to deal with different levels of complexity and variability in the data. The inherent complexity of any IR problem will always depend on the particular registration scenario considered. In general, two main determinant factors influence this complexity, that will be discussed as follows.

1.2.1 Technological means

In essence, the characteristics of the acquisition device, specifically the type of sensor and its underlying technology, determine the quality of the 3D images. Fortunately, thanks to the rise in the demand of 3D imaging content there are a wide variety of devices based on different scanning techniques: 3D laser scanners, structured light scanners, or time-of-flight (ToF) cameras. Along with 3D cameras, photogrammetry (a technique used to extract distances from 2D photographs and reconstruct 3D scenarios) is also used to produce range images representing the 3D structure of a scene. Besides, different types of 3D medical imaging techniques are also alternatives to generate volumetric 3D images. The main goal of this dissertation is to identify potential sources of complexity affecting the 3D IR procedure and address them using SC and CV techniques.

Image quality is an essential aspect of the IR process, particularly when considering three-dimensional data. We have identified potential scenarios with a considerable degree of complexity affecting the registration procedure as consequence of different degrees of quality. Therefore, the lack of technological means can be addressed from two perspectives, related to the quality of the images involved in the process.

- *Poor quality images acquired by low tier acquisition devices*

  Despite all the available options, in most situations it is not possible to employ a cutting-edge device. Budget, time, computing capacity, object scale or other characteristics of the scene are determinant factors that influence the accessibility of the artifacts, and therefore the quality of the images to be registered [BSH+13]. For example, when considering medical image applications (specifically radiation-based technologies such as CT, PET, or PET-CT

scans), the amount of radiation dose highly influences the quality of the resulting image [LKKB⁺15, Gol07], yet clinical protocols exist to protect the patient and restrict the radiation dose at acceptable levels [GM13].

Several IR problems involve low-quality imaging data or incomplete information. For instance, ToF cameras are a recently developed depth technology based on a simple principle: measuring the phase shift (time) between the illumination and the reflection of a light pulse, translating it to distance units, and generating a depth map (or point cloud) for every point in a scene. Compared to other devices, ToF cameras consist of promising features: low cost, portability, high frame rate (3D video recording) and very low computational requirements. However, the main shortcomings of ToF devices is their limited operating range, sensitivity to scene illumination conditions and low resolutions. Moreover, other limitations inherent to the ToF technology are the appearance of systematic errors related to the modulated light signal, resulting in distance and amplitude related errors due to interference with background light and object reflection. These errors introduce different amount of noise in the resulting range images by under or overexposing parts of the scene and miscalculations in the position of points in the 3D space. Hence, tackling IR applications where input images are acquired using ToF cameras entails a considerably complex problem.

- *High quality images along with large amount of data to be processed*

  3D medical imaging applications are one of the most relevant fields for basic processing tasks. IR is a prerequisite for further procedures in medical application, such as image segmentation, fusion or visualization, functional brain mapping, deformable atlas registration, or image guided surgery, among others. Accuracy and speed are the main factors to consider when evaluating the performance of any IR method [Gos12]. The general insight relates a higher image quality to more accurate or better registration performance, although that assumption is not rigorous [KDH⁺17]. It has been proven that lower quality implies noise, lower spatial resolution, or deformation/distortion of the subject, aspects that increase the complexity and influence the registration procedure [Zit03]. With higher quality, the complexity of the problem is also increased, as the amount of data can grow exponentially, exceeding the processing capabilities of modern machines. Therefore, the computational effort might become a limitation affecting the performance of the registration. Ultimately, the type or modality of the input image (which depends itself on the acquisition device) will be determinant in the performance of the IR method [GPM06].

1.2.2 Problem context/Scene factors

In other situations, the source of the complexity is not associated with the nature of the instruments or the quality/quantity of the data but with the specific scenario at hand which results in an inherently complex optimization task. Usually, a simple rigid transformation (involving rotation and translation) is considered in registration tasks that require reconstructing 3D objects from multiple range images acquired from different viewpoints. In controlled scenarios, ordinary IR methods exhibit a solid performance in normal conditions, unless the scene factors hinder the registration. Thus, when the misalignment between images is considerable or the overlapping areas of the object are minimal, ordinary IR methods are bound to provide a large misregistration. In addition, a specific application may require a more flexible transformation model by considering a shift in the scale of the objects, or accounting for distortions or deformation models, as is the case of many 3D medical imaging applications. Along with the increment in transformation flexibility, the search space is extended and therefore the optimization procedure becomes more of a challenge.

Additionally, tasks involving the registration of 3D models with 2D photographs pose another important and challenging problem. Usually, the transformation relating both images is a combination of a projective transformation and a rigid transformation used to determine the pose of the 3D model relative to the camera location in the photograph. Such is the approach followed by automated CFS applications based on computer-based skull-face overlay (SFO) techniques. Using SFO, a forensic anthropologist can determine if the model of the skull correspond to a specific person by projecting it over the other the 2D photograph and analyzing their morphologic characteristics. This procedure requires the anthropologist to consider correspondences between two sets of points, locating cephalometric landmarks in the face, and craniometric landmarks in the skull. SFO is a really complex optimization task, as it involves corresponding two different objects (a skull and a face) and the information leading the IR process is vague and incomplete, resulting in a huge search space with several local minima.

As a conclusion, solving the IR problem requires a careful examination of all the elements involved in the scenario under consideration (instruments, objects/subjects, scene context, expert knowledge, etc.). It should be acknowledged the extreme difficulty of designing a universal IR method. Traditional methods are usually limited to a reduced range of application. SC techniques are valuable tools capable of providing solutions where ordinary approaches are not able to. In particular, bio-inspired IR approaches are a suitable solution for facing this sort of complex optimization problems, allowing for versatility and scalability. Thanks to their global optimization nature, the search strategy of nature-inspired algorithms presents a robust performance when dealing with ill-defined IR problems [SCD11, DCS11].

1.3 Objectives

The main objective of this PhD dissertation is the development of new evolutionary IR methodologies in complex real-world scenarios from a multidisciplinary perspective. To that end, we have considered three different areas of application. For each potential application, the relevance of typical registration methods has been scrutinized in order to uncover the potential issues impacting their performance. Once the complexity factors are determined, we will undertake the following three lines of research in order to achieve the overall goal:

- Assess the viability of the IR problem involving low-quality input images. The objective comprises a study of the performance of ToF cameras compared to similar 3D acquisition devices and the development of a specific 3D reconstruction method for ToF range images. In particular, the design of an image preprocessing pipeline involving the enhancement of ToF range images, a subsequent feature extraction procedure to handle the complexity of the scenario, and the development of a novel bio-inspired IR method. This line of research is focused on dealing with complex scenarios caused by a lack of technological means in which the quality of the images is low.

- Address a complementary problem related to scenarios involving a lack of technological means, defined by the large amount of data contained in high-quality 3D medical images. This objective involves the extension of the methods developed for the previous objective to handle a 3D medical problem employing MRI techniques, which entails a high-quality registration problem. In this case, the approach requires a dimensionality reduction procedure, usually considering interest point detection methods, which tend to lose valuable information.

- Apply SC techniques to improve the capabilities of computer-based CFS approaches. Current automatic SFO methods project the 3D model of the skull (including the jaw as a rigid component) over the 2D ante-mortem photograph of the face. During the acquisition stage

of the 3D skull model, the pose of the mandible is usually approximated to the pose of the face in the photograph. Such a simplification causes a negative impact on the accuracy and reliability of the CFS technique, as this procedure is not anatomically correct and likely lead to missidentifications in certain conditions. In addition, one of the crucial criteria to evaluate the quality of a superimposition is the correspondence between teeth in the photograph and the skull [ICDS11, ICD12]. Hence, this objective involves modeling the mandible aperture movement and considering the jaw as a 3D articulated object independent of the skull. Several criteria must be taken into account when designing such a convoluted process, leading to a highly complex optimization task. Due to the novelty of this approach, the design of a ground-truth database will be required to assess and validate the reliability of the proposed registration method.

1.4   Structure

This PhD dissertation is divided into two parts. The first one is devoted to the statement of the IR problem and a detailed description of the different areas of applications identified for presenting a high level of complexity in addition to the presentation, analysis and discussions of the work developed. The second one collects the main scientific publications obtained as a result of the work developed.

In detail, Part I is organized as follows. First a statement includes an introduction to the problem addressed in this work, a justification establishing the foundation of this dissertation, and a detail of the proposed objectives. Then, Section 2 details the core of the IR problem and a revision of the current state of the art. Later, Section 3 introduces the particularities of the three complex real-world applications considered in this work: ToF range image registration (RIR), 3D Medical IR (MIR), and CFS, describing and analyzing the results obtained. Finally, Section 5 concludes with a discussion of the developments presented in this work and draws some future research lines.

Part II details the work developed to accomplish the objectives stated before by assembling the following three scientific publications:

- E. Bermejo, O. Cordón, S. Damas, J. Santamaría. Quality time-of- flight range imaging for feature-based registration using bacterial foraging, Applied Soft Computing, vol. 13, no. 6, pp. 3178–3189, 2013. DOI: 10.1016/j.asoc.2012.08.041. Impact factor (JCR 2012): 2.140. Category: COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE. Order: 23/114. Q1.

- E. Bermejo, O. Cordón, S. Damas, J. Santamaría. A comparative study on the application of advanced bacterial foraging models to image registration, Information Sciences, vol. 295, pp. 160-181, 2015. DOI: 10.1016/j.ins.2014.10.018. Impact Factor (JCR 2014): 3.893. Category: COMPUTER SCIENCE, INFORMATION SYSTEMS. Order: 8/135. Q1.

- E. Bermejo, C Campomanes-Álvarez, A. Valsecchi, O. Ibáñez, S Damas, O. Cordón. Genetic algorithms for skull-face overlay including mandible articulation, Information Sciences, vol. 420, pp. 200-217, 2017. Impact Factor (JCR 2016): 4.832. Category: COMPUTER SCIENCE, INFORMATION SYSTEMS. Order: 7/146. Q1.

2 Image Registration

This section provides a general introduction to the IR task and reviews the literature analyzing both traditional and sophisticated approaches to address the challenges of this problem.

2.1 Problem Definition

Independently of the application, the registration framework is defined as an optimization problem that aims to maximize the similarity between aligned images. Usually, the process involves two input images located in different coordinate systems: a reference or *model* image (I_M), and a *scene* image (I_S), which is transformed to reach the model geometry. Thereby, the goal is to find a geometric transformation T that applied over the scene $T(I_S)$ maximizes its similarity with the model I_M. The quality of the alignment between images is measured by a similarity metric F. Formally, IR can be represented as a maximization problem over the space of transformations:

$$\underset{T \in \text{ Transformations}}{\text{argmax}} \quad F\left(I_M, T(I_S)\right).$$

Even though there is no universal IR method that could be suitable for every circumstance, usually a common framework is defined regardless the specific application. Hence, an iterative optimization procedure is frequently considered (see Figure 1), where an IR method consists of a combination of four components: i) a *feature space*, which extracts the intrinsic information of the images; ii) a *search space*, defining the spatial transformation model used to align the images; iii) a *similarity metric*, which measures how similar are the images after the registration; iv) a *search strategy* or optimizer, responsible for finding and refining the solution until the optimal transformation is found or a particular stopping criteria is reached (i.e. when the algorithm performs a determined number of iterations). The said components are further detailed in the following paragraphs.
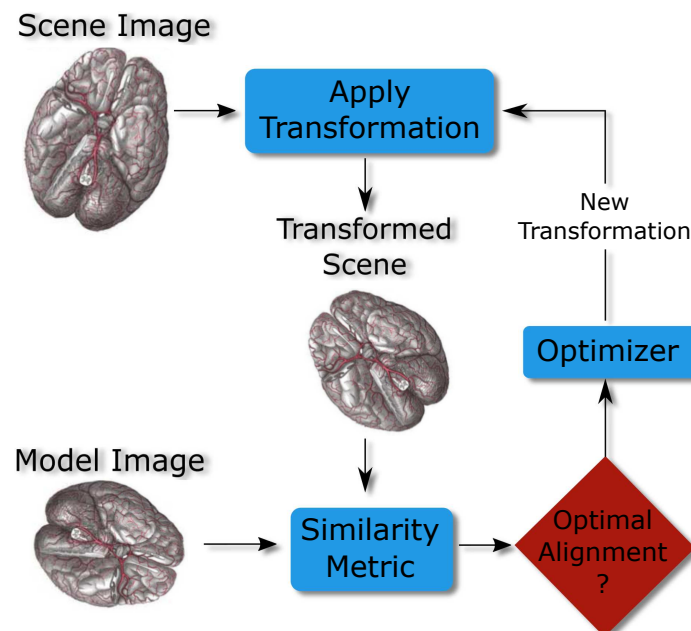

Figure 1: The interactions among the three components (blue boxes) of a registration technique.

2.1.1   IR components

Feature space: image information

Depending on the nature of the information used during the registration process, IR methods can be classified in two main categories: *intensity-based* or *feature-based*. The former directly use the original data without applying any preprocessing technique. In particular, the intensity patterns of each image are used to measure the similarity between images via mathematical or statistical criteria. These methods are based on the premise that the optimal registration will maximize the similarity. Even though intensity-based approaches are frequently more accurate due to dealing with the totality of the information contained in the scene, this procedure is extremely demanding in computational terms. Therefore, sampling procedures are usually applied to improve the efficiency of the methods. An alternative is to consider intensity-based methods as a fine-tuning registration process. Their main disadvantage lies in the similarity measurement procedure (usually rectangular or circular windows are used to establish a relationship between images [Gos12]). Since similarity measurements take place among corresponding parts of the images, when considering complex deformable transformations the corruption of the spatial intensity relationship leads to failures in the measurements during the registration procedure [MR13].

Feature-based approaches aim to deal with the downsides of intensity-based methods by only considering critical information about the scene [Roh01, WEF$^+$08]. In order to achieve a status of reliability, the feature detector must satisfy the requirements of independence and distinction to spatial neighbors to avoid redundant information, computational efficiency, comparability, invariance, and robustness against noise, blurring, contrast, radiometric conditions, and luminance or geometric changes [Zit03, FBC13]. Usually, feature points are identifiable or prominent elements of the images, such as points [OCL06], curves [MV98], and regions [FS94]. The extracted feature points are then used as image descriptors to represent the relevant areas of the images and facilitate the estimation of the similarity measurement by decreasing the amount of data. Usually, feature-based methods achieve a coarse approximation to the global solution and a following refinement stage is applied in order to improve the accuracy of the registration.

Another factors to categorize IR methods are: i) the dimensionality of the images: *2D*, *3D* or a combination *3D/2D*, ii) the number of devices used during the acquisition of the images: *monomodal* (same device) or *multimodal* (different devices), and iii) the nature of the scene: *intrasubject* (when the images belong to a single object or subject), *intersubject* (when the images are obtained from multiple but similar objects/subjects), or *atlas-based* (when one image belongs to a single object/subject and the others come from an information database [MR13]).

Search space: transformation model

The structure and extent of the search space rely on the transformation model considered by the IR method, as it determines the characteristics of the geometrical transformation used to align the images. Again, there are two main categories, depending on the nature of the transformation model: *rigid* (linear/global) and *non-rigid* (elastic/local). Application context determines the proper model to be used depending on the geometrical properties (e.g., orientation, shape, size, etc.) required to be preserved. Linear models (i.e., rigid, similarity, affine or projective transformations) alter the entire images while preserving distances, lines and angles. In contrast, elastic transformations (radial basis functions such as B- or thin-plate splines, or diffeomorphisms) are used when the application requires local deformation of the images (allowing a local warping of the image features).

Mathematically, 3D linear transformations are represented by 4x4 matrices in homogeneous coordinates (as in Eq. I.1). Below, the characteristics of different linear transformation models are described.

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & w_{13} & t_x \\ w_{21} & w_{22} & w_{23} & t_y \\ w_{31} & w_{32} & w_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \qquad (I.1)$$

- **Rigid transformation:** Rigid transforms can be specified as a combination of a rotation operation per axis (R_x, R_y, R_z) and a translation T. Since the rotation operations are orthogonal, the angles and length of the original image are preserved.

$$T_r = R_x \cdot R_y \cdot R_z \cdot T,$$

where

$$R_x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad R_y = \begin{pmatrix} \cos\beta & 0 & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_z = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 & 0 \\ -\sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad T = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- **Similarity transformation:** It allows for a change in scale to adjust the aspect ratio of the objects while only preserving the angles, adding the corresponding scaling matrix S':

$$T_s = S' \cdot T_r,$$

where

$$S' = \begin{pmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- **Affine transformation:** By considering a shearing operation, affine transforms preserve the parallelization and intersection properties of the lines, but not their length or the angles in the original image.

$$T_a = H \cdot S \cdot T_r,$$

where

$$H = \begin{pmatrix} 1 & h_{xy} & h_{xz} & 0 \\ h_{yx} & 1 & h_{yz} & 0 \\ h_{zx} & h_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad S = \begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- **Perspective projection:** Projective transforms are usually considered when dealing with 3D/2D registration applications to apply a perspective distortion mapping the 3D model to the 2D photograph. This type of transformation does not preserve the line properties when mapping lines to lines. When the coordinates of the objects in the 3D scene are known (x_0, y_0, z_0), the corresponding 2D points in the photograph are defined as a function of the optical lens center:

$$x_i = \frac{-f x_0}{z_0 - f}, \quad y_i = \frac{-f y_0}{z_0 - f}$$

Usually, IR applications considering range images aim for the reconstruction of real-world scenarios [BL95, SBBG03, ATS+03, BLR03]. As this kind of images are acquired directly from the scene, the objects are represented in the same scale space and IR methods employ rigid transformations. Similarly, most medical applications involving anatomical structures like bones or the brain [LYJ03, ACL04, HPO+05] consider rigid transformations to avoid unrealistic representations of the skeleton or tissue growth.

If scaling is also required, a similarity transformation is then considered. In contrast, applications using CT scans may require non-uniform scaling due to the presence of shearing. In these cases, the considered transformations are affine. When soft tissue is involved and deformation can be expected, deformable transformations are commonly used [BT01, Mey07, ZMS+10].

Similarity metric: alignment measure

The similarity metric, noted as $F(I_M, T(I_S))$, assesses the quality of a transformation T in terms of the degree of resemblance between the model I_M and the transformed scene $T(I_S)$. In other words, it measures the similarity between images after applying the estimated transformation and provides a way to compare the result of the registration. In order to provide an informative comparison, the measurement should be robust to noise and background changes. Therefore, a careful choice of the similarity metric is a decisive step in the design of any IR method [SMA76]. In particular, this choice will depend on the nature of the registration approach. Hence, similarity metrics are divided in two distinct categories whether they rely on image attributes or image geometry:

- *Feature-based* approaches deal with coordinate primitives, i.e., the points extracted by the feature descriptor to measure the similarity. Hence, the properties of the feature points are a determinant factor in the performance of the registration. Feature-based methods usually consider mean square error (MSE) [AFP00, MRDG02, ATS+03, CS05] or median square error (MedSE) [SCDI09, RFL02]. To evaluate the alignment, each point x_i in the model is paired with the closest point in the transformed scene, computing the MSE as:

$$\text{MSE} = \frac{1}{r} \sum_{i=1}^{r} \|x_m - c(T(x_s))\|^2, \tag{I.2}$$

  where r is the number of points and c is the correspondence criterion establishing a matching between pairs of model and transformed scene image points. Meanwhile, MedSE is defined as the median value of the Euclidean distance between corresponding points:

$$\text{MedSE} = median(\|x_m - c(T(x_s))\|^2). \tag{I.3}$$

- In *intensity-based* approaches, elemental primitives (intensity values) of the two images are considered and the relationship between intensity distributions is measured. Common intensity-based metrics are sum of squared differences (SSD) [OSP02, DM03], normalized correlation (NC) [Lew95, Son11], and mutual information [CMD+95, VW97, PMV03]. Their suitability is determined by the relationship between images and the employed acquisition technique. SSD can be used when images present the same gray level structure, so it can be applied for monomodal registration under the assumption that images differ only by Gaussian noise.

  If this structure is non existent but a dependency between gray level is supposed, NC is a viable alternative assuming a linear relationship between the images. Meanwhile, in those cases where images are acquired using different sensors (multimodal registration), the relationship can be strongly non-linear and metrics based on information theory are the most suitable option.

The computational cost of the similarity metric is often considerably large due to the huge amount of data required by the IR process, resulting in a complex and time-consuming task. In order to improve the efficiency of the optimization process, most IR contributions define a spatial indexing data structure. This data structure is built once at the beginning of the procedure to expedite the metric calculations allowing for a fast retrieval of the closest point assignment. In the specialized IR literature, two main data structures stand out:

- *Kd-tree*, which is a data structure used to organize points in a space with k dimensions. It is a binary search tree with different constrains, corresponding to the k dimension generalization of the 1D bisection. The tree is decomposed in the different dimensions while maintaining the neighborhood relationship in within its structure.

- *Distance map*, where each cell of the structure stores the Euclidean distance to the closest point of the corresponding image. A particular case of distance map is the grid closest point (GCP) [YAF99], which consists of two cubes splitting the 3D space.

Search strategy: optimization procedure

This component refers to the strategy followed to explore the search space of the geometrical transformation. It consists of an iterative optimization process designed to find the optimal solution within the space of spatial transformations in terms of the similarity metric. Many factors are accountable for the complexity of an IR method, as altogether define the topology of the search space. Since usually the search space is highly multidimensional with several local optima, exhaustive search strategies are impractical approaches. Exact methods following this approach are not able to find an accurate solution efficiently unless the addressed problem is extremely simple. In essence, a common approach is to establish a compromise between performing an exhaustive search and accepting a local optimum in order to deal with the non-linearity of the IR optimization problem. Following this alternative, approximate methods implement a heuristic optimization approach (e.g., greedy, local search, or hill climbing methods) to improve the exploration strategy and reach to better solutions.

According to the characteristics of the search space, we can find two different strategies to determine the parametric transformation. When the search is carried out in the space of the transformation parameters, the registration is treated as a continuous optimization problem by *parameter-based* IR approaches. Once the optimizer estimates a set of transformations parameters, the validation of a particular solution is computed by matching the model and scene considered primitives. Alternatively, the space of feature correspondences is an alternative way to perform the search. Thus, *matching-based* IR approaches are able to compare either areas of the image (intensity-based methods) or matching features (feature-based methods). Transformation parameters used to align the images are derived after the matching of model and scene features is accomplished.

Two different techniques are widely used to improve the performance of the search strategy following a coarse-to-fine approach. The first technique considers a two-stage registration procedure. An initial prealignment stage (based on a global optimization strategy) provides coarse results close to the optimum. It is followed by a refinement stage to adjust the results, usually considering a method based on a local optimizer as Iterative Closest Point [BM92] (see Section 2.2.1). The second technique is known as multi-resolution or pyramidal strategy. The registration starts the iterative process using a low resolution (downscaled and/or smoothed) version of the input images. In this stage, the computation of the similarity metric is faster due to a reduction in the amount of data while the search space is smoothed (presenting a lower amount of local minimum) as a result of image details becoming imperceptible. Once the optimizer finds a solution, the procedure enters its next stage. Here, the optimizer adapts the previous solution while dealing with more detailed image data.

2.2   State of the Art in Image Registration

This Section introduces a general overview of the literature from classical optimization methods tackling the IR problem to novel evolutionary computation-based proposals. In addition, a

justification of the suitability of MH-based approaches as robust and efficient alternatives is provided for those problems where the complexity of the problem is not trivial.

During the last few decades, several surveys describing different taxonomies have been presented in an attempt to classify the wide variety of IR methods proposed so far. The first classification was published by Brown (1992) [Bro92], followed by Zitová and Flusher (2003) [Zit03], Modersitzki (2003) [Mod03], and Goshtasby (2005) [Gos05], establishing a well-defined organization of IR approaches from a general point of view. These taxonomies consider different criteria: image dimensionality, image acquisition procedure, nature of the registration basis, domain of transformations, search strategy, field of application, etc. Meanwhile other classification proposals limit the review to a specific criterion, such as field of application (remote sensing [FM96], range imaging [SMFF07], or medical imaging [MV98]), similarity metric [MRH10], transformation model [SDP13, TCL$^+$13], or a combination of some of them [PMV03, DCS11]. An extensive survey of the matter is beyond the scope of this dissertation, so the analysis will focus on a particular criterion to review the relevance of IR contributions in the literature.

2.2.1 Classical Optimization in Image Registration

We begin the analysis by establishing the search strategy as the criterion that will help us organize and differentiate traditional IR proposals. It was noted earlier that the optimizer performs the search either following a *parameter-based* or a *matching-based* approach. Classic numerical optimization algorithms stand out when facing the IR problem under a *parameter-based* approach. In order to perform the search, each solution to the IR problem is encoded as a vector containing the values of the parameterized transformation f. The similarity metric F is responsible for guiding the search, which evaluates the solution vector resulting in a numerical optimization problem.

As stated before, the characteristics of the search space are highly dependent on the considered transformation model. Nevertheless, the orders of magnitude in the scale of the parameters are an essential issue to consider. In terms of displacement, the impact of unit changes in rotational parameters are far greater than translation parameters. The scale shift between elements in the solution vector results in elongated valleys within the search space.

Among classical optimization methods we can find Newton's methods, Levenberg-Marquardt [Lev44], gradient descent, conjugate gradient descent [HS52], Powell's method [Pow64], Nelder-Mead Simplex [JR65], Solis and Wets method [SW81], etc. Numerical methods face many disadvantages, especially considering the local-search nature of the optimization. A large amount of misalignment between images, the presence of noise, and the aforementioned difference in scale of the registration parameters are significant difficulties for these optimizers [BM92, HN02]. Moreover, they present inherent problems regarding the estimation of derivative information when dealing with non-linear similarity metrics.

Some contributions have analyzed and compared the performance of numerical optimization methods within the framework of medical IR [MVS99, KSP07]. Two main conclusions were drawn in these comparisons. First, the importance of multi-resolution strategies within the optimization procedure to improve the convergence speed of the IR methods, even though its application is not sufficient to ensure the location of the global optimum. Second, the dominance of stochastic gradient descent, in particular the Robbins-Monro method [RM51], which obtained the best performance in different applications. As a result of these studies, Klein et al. [KPSV09] introduced an improved version of the Robbins-Monro method specifically designed for IR, known as adaptive stochastic gradient descent. Despite the difficulties of numerical optimization methods to avoid local minima, approaches integrated in global schemes (specifically designed to address a particular problem) have proven their reliability and effectiveness [JS01].

Regarding *matching-based* approaches, it is worth noting that the optimization procedure is not guided by the similarity metric but rather by the computed matching between pairs of

corresponding image primitives. Once the correspondences are found, direct methods can extract the parameters of the transformation using a least-squares estimator [AHB87], m-estimator [Hub81], or other model fitting techniques [ZSML10]. It is possible to obtain an accurate and efficient estimation of the parameters from at least three correspondences, even though in practice, a broader set of observations is required to ensure it [Hor87]. This requirement is motivated by the presence of outlier correspondences that might alter the correct estimation, mainly due to the presence of noise in the input images.

Many matching-based methods have been proposed to deal with the spatial relations among primitives, e.g., graph matching [Gos12], clustering techniques [SG87], or chamfer matching algorithms [BTBW77], but one approach stands out in the literature. The Iterative Closest Point (ICP) [BM92] is by far the most relevant traditional method, widely used in range IR applications. ICP is an iterative procedure that applies a transformation to the current position of the scene image at each iteration. This transformation displaces the image primitives in a least squares sense as close as possible to their corresponding primitive in the model. This procedure involves two stages. The first one requires the identification of the corresponding model point for each point in the scene. Whatever the original representation of both sets of image primitives are, they are converted to sets of points in order to compute their correspondences. Then, every scene point p_i is taken into account to search for its closest point in the model image m_i so that the distance between correspondences is minimized. This process usually considers MSE (see eq. I.2) as similarity metric. Meanwhile, the second stage involves the estimation of the rigid body transformation relating these points using a least-squares estimator. The iterative procedure comes to an end if the change in the error measurement falls below a determined threshold τ.

ICP is able to obtain accurate results even in presence of Gaussian noise. However, three main drawbacks have been identified:

- ICP cannot cope with non-overlapping regions. As outliers are never removed and influence the procedure, one of the images (preferably the scene) has to be contained in the other.

- The convergence of the algorithm to a global optimum is not guaranteed, as the algorithm always converges monotonically to a local minimum. ICP is very dependent on the initial guess and it is prone to get trapped in local minima, so a prior alignment of the images is required to ensure convergence to a correct solution.

- ICP can only handle normally distributed sets of primitives. If the data set is non-uniform primitive information is not preserved across the image, the appearance of discontinuities or implicit ambiguities will negatively affect the computation of the correspondences.

Different modifications have been proposed to overcome these drawbacks while improving the robustness to outliers and computational efficiency: introducing correspondences to tangent planes in the model [CM92], considering kd-trees to accelerate the search of correspondences [Zha94], considering invariant features to avoid local minima [SLW02], or using collinearity and closeness constraints to approximate the neighboring region of the points [LSD04], among many others.

2.2.2 Metaheuristics and Evolutionary Computation in Image Registration

With the limitations of traditional IR methods and the increasing number of contributions introducing adjustments to expand their applicability towards global optimization techniques, basic heuristic and MH approaches [Glo03] became an appealing alternative to deal with complex IR problems [YAF99, RJR00, CEGL01, CDS06b]. In particular, MH approaches are among the most prominent interdisciplinary techniques when considering numerical optimization and computationally hard combinatorial problems.

MHs are approximate and non-deterministic strategies guiding the search procedure with the ability to efficiently explore large search spaces and easily adapt to a wide range of problems. As opposed to numerical optimization methods, MHs are able to obtain high quality solutions in problems with challenging space topologies, though relaxing the guarantee of finding the optimal solution. So far, many MHs have been proposed and compared under different criteria [BR03], [BDGG09]. These approaches are usually differentiated regarding: the origin of the algorithm (nature or non-nature inspired), the number of solutions used at the same time (single search/trajectory or population-based), the objective function type (dynamic or static), the number of neighborhood structures (one or many), and the use of search history (memory or memory-less).

Through combination of the previous characteristics many MH approaches have been proposed. Iterated local search (ILS) [LMS03], variable neighborhood search [MH97], simulated annealing (SA) [KGV83], memetic algorithms [Mos89], greedy randomized adaptive search procedures (GRASP) [FR89], tabu search [Glo89], and scatter search [Glo77, LM03] stand out as the most renowned ones. In addition, some subfamilies of MH algorithms gained notoriety due to their efficient performance in many real-world problems. Such is the case of swarm intelligence (stochastic diffusion search (SDS) [Bis89], particle swarm optimization (PSO) [KE95], or ant colony optimization (ACO) [DS04]), and evolutionary computation (evolutionary algorithms (EAs) [Fog06] such as genetic algorithms (GAs) [Hol92, Gol89], genetic programming [Koz92], evolution strategies [BD01], evolutionary programming [BFM97], and differential evolution (DE) [SP97]).

The approximation used by EC methods is to use computational models of evolutionary processes found in nature to evolve populations of solutions (individuals) iteratively, inspired by the Darwinian theory of evolution and natural selection. Each iteration corresponds to a generation where different modifications are introduced stochastically in the population to create new and better individuals. The main idea behind this analogy is to design a versatile optimization algorithm able to maintain a balance between search space exploration and exploitation of the information related to good solutions. This trade-off is extremely important, as it allows the algorithm to rapidly identify search space regions with optimal solutions improving the efficiency and avoids previously analyzed or low-quality regions. Particularly, EAs are probably the most used evolutionary approach to deal with real-world optimization problems.

In the last few years, the success both in performance and efficiency terms of swarm intelligence (SI) [KES01] and evolutionary methods both in terms of performance and efficiency when applied to real-world problems have significantly increased the number of contributions in the field. In particular, nature-inspired (NI) computing has gained a large attention from researchers and has become an expanding area of interest [FYBF13, SA15] but it has also stir some controversy [Sör15]. Several methods have been proposed following the allegory of natural phenomena [Yan14], such as artificial bee colony (ABC) [JJCJ16], cuckoo search [YD10], firefly algorithms [Yan09], or bacterial foraging optimization algorithm (BFOA) [Pas02].

The first application of EC to an IR problem can be traced back to the eighties, where an approach based on a GA was proposed to align 2D angiographic images [FGVg84]. Since this initial contribution, evolutionary IR has become an incredibly active and successful research field. Several proposals followed this trend of applying EAs to different IR applications, being GAs the most prolific approaches. Mandava et al. [MFP89] performed a search space scaling for 2D calibrated clinical images. Brunnstörm and Stoddart [BS96] introduced a two-stage IR approach where a genetic-based coarse prealignment tackles the free-form surface matching problem. Tsang et al. [Tsa97] employed binary-coded solutions for 2D shape alignment of different objects. Yamany et al. [YAF99] included a GCP structure and considered rigid transformations to tackle 2D and 3D surface registration. Chalermwat et at. [CEGL01] considered a multi-resolution strategy and binary-coded solutions to register wavelet compressed images.

Several pitfalls affect the latter approaches, specifically discretization flaws (unexplored regions of the search space) as a result of the consideration of binary coding to solve problems whose parameters belong to continuous domains. Real-coded GAs were used by Rouet et al. [RJR00] for rigid and elastic 3D medical IR. Lomonosov et al. [LCE06] compromised the precision of real-coded solutions by considering integer-coded representations to obtain a faster convergence of the GA. Other MHs have also been applied to IR. Salomon et al. [SPH01] considered a DE algorithm to tackle 3D medical IR. Cordón et al. proposed in [CDB03] an approach based in ILS for the registration of 2D image contours while in [CDS03] used a CHC EA [Esh91] for 3D synthetic images and MRI IR. Maglogiannis [MZ04] proposed a SA method to deal with the automated registration of several types of 2D medical images.

The suitability of MH and NI approaches have been meticulously analyzed stating several strengths and limitations that serve as arguments to justify or avoid certain complex optimization problems such as IR [OCL06, NKDV09]. The most important advantages related to IR are:

- Most MHs are conceptually simple and easy to implement. They are more robust than traditional methods, as they do not depend on the starting solution (as opposed to ICP or gradient-based methods). In addition, specific strategies are usually provided to escape from local optima [YAF99] and deal with multimodal functions to address IR problems [PT07].

- EC and MHs demonstrate versatility and efficiency in a wide variety of optimization tasks within IR, dealing with combinatorial problems such as facing both the transformation parameters and the matching-based IR approaches.

- The optimization scheduler of MH approaches is easily adaptable to specific constraint requirements and objectives. For example, EAs can be integrated in a multi-objective scheme to tackle IR problems [SBB05].

- In contrast to other numerical IR techniques only suitable for continuous functions or constrained sets, EAs are independent of the solution representation performance-wise.

- It is easy to include prior knowledge about the problem to their design, yielding a more efficient exploitation of the search space. For example, feature-based approaches are able to exploit the primitive information (geometry and/or intensity) of the images [BS96, CDS06b, CDSM08].

- MHs can also be easily combined (e.g. hybridization) with traditional optimization techniques to refine the exploration-exploitation trade-off [SCDI09, XD04]. The stochastic strategy is usually followed by a deterministic (gradient-based) method.

Despite the virtues of EC and MH algorithms, they are not exempt of important weaknesses:

- Most NI approaches require a manual, error-prone tuning phase of control parameters.

- Usually, MHs are time-consuming processes requiring many iterations, making them unfeasible alternatives when real-time is required.

- Many MHs lack a formal proof to assure optimum convergence.

- The estimation of a suitable stop criterion is not trivial. It depends on problem context, although CPU time or number of function evaluations are common choices.

In order to overcome these shortcomings, some proposals include advanced strategies specifically designed for IR, such as adaptive control parameters [OLZW06, SDGTC12, SDCE13], automatic parameter tuning [VDLS+13], parallel or GPU implementations [RF02], or hybrid approaches to escape from local optima [SCDI09]. In the next section, the most relevant evolutionary IR methods are briefly described.

2.2.3   State of the Art Image Registration Metaheuristics

Given the high amount of applications and the wide variety of IR methods specifically designed for a particular problem, many classifications of relevant methods in the literature can be analyzed. The distinctive characteristics of the problems addressed in this PhD dissertation narrows the analysis to methods following a feature-based approach where the optimization is performed in the space of transformation parameters. Precisely, the following methods are considered the state of the art in evolutionary IR and will serve as comparison for the approaches proposed in this work:

- **Luck et al.'s ICP-L**
  The authors proposed in [LLH00] a hybrid IR strategy combining two stages: i) First, the ICP algorithm is applied while discarding outlier correspondences which distances exceed a determined threshold, ii) Then, a SA algorithm refines the generated solution following a parameter-based approach with a reduced search space exploration. The resulting algorithm is a robust approach able to find an optimal solution in a reduced number of iterations. It was originally applied to the range IR problem.

- **Chow et al.'s GA-C**
  [CTL04] consists of a parameter-based genetic algorithm considering 3D rigid transformations and six-dimensional real-coded chromosomes (solutions). The first three elements of a solution correspond to the rotation component in terms of euler angles, while the rest correspond to the translation component. The novelty of this proposal resides in the integration of two novel operators: a crossover that randomly selects the number of genes to be swapped, and a mutation that randomly selects the number of genes to be swapped. The value of a rotation gene is generated randomly within a constant range and the translation gene is dynamically computed according to the fitness value of the chromosome. These operators allow the GA to reduce the magnitude of the mutated parameters when the individual is close to the optimum. The algorithm also includes a sophisticated restart mechanism known as dynamic boundary. Once the search reaches a stationary state and the fitness of the population stall the improvement, the search range is concentrated to the area of the best individuals by restricting the valid ranges of the genes in a chromosome.

- **Silva et al.'s GA-S**
  In [SBB05], the authors tackled a range IR problem with images acquired from 3D laser scanners considering 3D rigid transformations. The considered problem was challenging due to the large amount of misalignment between images and low overlapping. The proposed method was based on steady-state GAs [BFM97] where 90% of the worst solutions are replaced instead of following a generational scheme. The GA followed a hybrid scheme considering tournament selection, uniform crossover, and random mutation, along with a hill climbing algorithm as solution refinement. In addition, the authors proposed a novel similarity metric, surface interpenetration measure improving the accuracy of the results provided by other Euclidean distances.

- **He an Narayana GA-HE**
  The method proposed in [HN02] is another hybridization where a real-coded GA is combined with a deterministic global optimization method based in branch and bound and called DIRECT [JPS93]. Their approach follows a two-stage approach where the GA is applied as a coarse registration stage prior to a refinement of the solution using DIRECT in a reduced search space. The authors evaluated their proposal in a 3D medical registration problem considering MRIs and 3D rigid transformations.

- **De Falco et al.'s DE-EA**
  The authors proposed a method based in DE [DDMT08] to tackle two different 2D IR problems: changes in time of satellite images and mosaicking. DE combines a set of arithmetic operations with the classical crossover, mutation, and selection operators of a GA. This method has the advantage of considering a small amount of control parameters (mutation factor and recombination rate). Solutions are modified using a weighted difference of two different individuals in the population. The proposed DE method followed a parameter-based approach considering a mutual information similarity metric and affine transformation models with seven-dimensional real-coded solutions.

- **Wachowiak et al.'s PSO-W**
  The method proposed in [WSZ+04] is based on PSO and focused on the registration of multimodal 2D and 3D medical registration problems. PSO is another population-based MH that exploits the cooperative and social behavior of organisms in bird flocks, insect swarms or fish schools. Every iteration, the individuals (particles) in a population (swarm) change their location in the search space guided by a velocity vector. The authors introduced different alternative designs of the original PSO scheme following a parameter-based approach, mutual information as similarity metric and rigid transformations. The main disadvantage of this proposal is its dependence on the initial orientation of the images before applying the registration procedure. Furthermore, the algorithm is designed to exploit an initial solution provided by the user, limiting the robustness of the proposal.

- **Santamaria et al.'s SS-S**
  In [SCD+07], the authors introduced a SS-based IR method following a parameter-based approach to 3D reconstruction of forensic objects. Seven-dimensional real-coded solutions were considered to represent rigid transformations along with a GCP structure to improve the efficiency of the algorithm. In [SCDI09] a broad study was performed comparing different MHs in a memetic scheme including a local search procedure to exploit the global capabilities of EAs. In particular the comparison included different IR methods: SS [CDS06a], CHC [CDS06b], DE [DDMT08] in combination with three different deterministic algorithms as local search: XLS [BD01], Solis and Wets [SW81], and Powell [Pow64]. The result of this comparison determined that SS in combination with XLS outperformed every other evolutionary approach in a memetic scheme. The authors evaluated their proposal in a IR problem tackling the 3D registration of different human skull models. In [DCS11] another exhaustive comparison of evolutionary IR methods was performed in a 3D medical IR problem. The SS memetic approach proposed in [SCDI09] confirmed the good results obtained previously and consolidated its dominance as state-of-the-art evolutionary IR method.

3   Complexity in Image Registration Applications

This section introduces the particularities of the applications addressed in this PhD dissertation. Two main issues have been identified as determinant factors for the complexity in an IR task: the lack of technological means and certain scene factors resulting in a complex problem from the optimization viewpoint. Different procedures can be applied to overcome these emerging challenges. In order to provide a comprehensive overview of this multidisciplinary contribution, the analysis is partitioned in three subsections. The first subsection introduces the novel developments in range imaging, the proposed pipeline to enhance the image quality of TOF cameras, and a novel evolutionary method for RIR applications. Next, the proposed evolutionary IR method is evaluated in a scenario involving a similar yet distinct problem whose complexity requires a different approximation: MIR. Finally, the third subsection details the problem of SFO and introduces the articulated SFO method proposed to extend and improve the functionality of the current approach.

3.1   Range Image Registration

Building 3D representations of ordinary objects in the real world is an extremely relevant topic in CV. 3D reconstruction, modeling, and recognition play a crucial role in many contemporary applications. Image acquisition and registration are essential and indispensable components integrated in these tasks. The process starts by sampling a real world scenario and capturing 3D images using a laser or structured light scanner, a stereo vision system, a ToF device, or any of the currently available techniques using 2D images, e.g., photogrammetry. The first obstacle usually appears in this stage as most acquisition devices are limited to a field of view and the captured scene appears partially represented in the images, presenting occlusion or blind areas. As a result, a number of images have to be acquired from different viewpoints to ensure a complete representation. A registration stage followed by a fusion procedure have to be applied in order to align the different views and reconstruct the entire model. Some modern devices rely on different calibrating mechanics and include a commercial software suite in an attempt to ease and automate the acquisition procedure to the final user.

Usually, the registration methods integrated in market applications are based on the well-known ICP algorithm, resulting in a procedure dependent on object pose, device location within a delimited working area, and accurate calibration. Therefore, the automatic registration commonly fails and the user is required to manually pre-align the different partial models. In addition, many applications involve complex scenarios where the object moves or consider large scale objects (buildings or archaeological sites). On these grounds, range IR is still considered an active and prolific research field, relevant in the development of several applications, such as reverse engineering, cast fabrication, or quality inspection in Manufacturing Engineering [LPGL16], landscape survey and artifact preservation, reproduction, and visualization in Archaeology and Cultural Heritage [LRG+00, SBBG03, ATS+03], assisted surgery in Healthcare [MRDG02, MBWH11], or face detection in Biometrics [NC09], among others.

3.1.1   Modern 3D Acquisition Technology

Several contributions analyze and compare the technology behind a wide variety of modern 3D imaging devices [STD09, GTKK13, PLB12], while others provide a generic overview of their principles of operation [SAB+07, HE11]. The main characteristics of available technologies are described below to provide a basic notion in the subject. Essentially, 3D imaging sensors are classified by the underlying principle used to measure distances in a 3D space. There are two main technologies: triangulation and time-of-flight.

- **Triangulation**: This technique follows a geometrical approach based on forming a triangle between a target point and two components of the measurement system (which is essentially the basis of human depth perception). As the position of the two measurement sensors is known, the distance to the target point is easily determined by epipolar geometry. There are two categories of triangulation devices:

    - *Passive triangulation* employs two receiver sensors that observe the scene from different locations to extract 3D information. This is the principle behind *stereo vision*. Simplicity and low cost are the main advantages of stereo vision systems. However, the quality of the images is highly sensitive to illumination, surface reflectance, and sharpness of the surface textures. As the computation of depth information is based on details and features, the performance of passive devices drops in scenarios with uniform or non-textured surfaces. In addition, a precise calibration is required in order to provide accurate measurements, which are affected by the correspondence problem. The location of corresponding features between the images is a complex and time-consuming process.

    - *Active triangulation* consists of an emitter projecting a light beam onto the scene and a receiver measuring the incidence angle of the reflected light. Two different kinds of devices employ this technique. *Structured light* devices ([GIV10]) are able to project 2D patterns of non-coherent light onto the objects and simultaneously generate 3D information. The pattern is distorted by the surface of the objects in the scene and the receiver is able to infer depth information from the deformation of the original pattern [SFPL10]. Structured light scanners are faster than stereo vision systems and, in the last few years, have evolved into accurate, versatile, and portable devices suitable for close-range applications. These devices also suffer from the correspondence problem, even though the use of patterns reduce the computational burden. Meanwhile, *laser scanners* ([HL09]) use a laser beam to sweep the scene either emitting a laser dot or a laser stripe to speed up the process while a receiver sensor (usually a position sensitive detector [CS10] or a mirror galvanometer [AE94]) measures depth information. Modern hand-held laser scanners are highly accurate devices equipped with a wide variety of sensors: internal or external trackers (to determine a reference system), infrared diodes (to provide resilience to illumination), and additional cameras (to enhance flexibility and capture colored textures). Their main advantages are accuracy and high resolution.

- **Time-of-flight**: ToF technology [RS13] is based on the same principle as radar. It measures the time needed for a specific light source to travel from an emitter to an object and back to the receiver. In this case, the receiver is usually a photosensitive diode. Two different methods can be applied: impulse time-of-flight and phase difference. The first approach directly measures time by assuming the signal speed as a known variable c. Then, the distance can be measured as: $d = c \cdot t/2$. In contrast, the phase difference method is based on the emission of a signal modulated in a determined frequency, e.g. 30 MHz. The difference between the emitted and the reflected lights produces a phase delay used to obtain the measurements:

$$D_i = \frac{c\varphi_i}{4\pi f_{mod}} \tag{I.4}$$

where f_{mod} is the modulated signal frequency and φ_i is the phase shift of the signal. This procedure allows the device to obtain thousands of measurements per second. In addition, multiple light pulses are used to improve reliability.

The main advantage of ToF cameras is their simplicity and efficiency. As opposed to laser scanners, there is no need for mechanical moving parts in ToF devices. Moreover, this kind of depth scanner is ideal for real-time applications as it can capture 3D scenes up to 160 frames

per second. In contrast, ToF sensors face important limitations. The most relevant is the low spatial resolution (commonly less than 1 Megapixels) of the CCD or CMOS sensors used as receivers, far away from the resolution achieved in consumer DSLR cameras (around 20-50 MP). Another drawback of ToF cameras is their low signal-to-noise ratio, resulting in several measurement errors due to the limitations of the emitted signal power. Accuracy is sacrificed for the sake of higher frame rates.

Furthermore, different systematic errors related to the sensor and non-systematic errors related to the scene context affect the performance of ToF devices. Among systematic issues, distance-related errors result from imperfections in the generation of a sinusoidal modulated signal while amplitude-related errors are the result of a low strength of the reflected signal, saturation or object reflectivity. Non-systematic measurement errors usually occur when ToF cameras illuminate the whole scene at once. As a result, sensor pixels have a high dynamic range leading to wrong measurements due to motion blur in the presence of dynamic objects, or scattering and interferences when the signal is reflected in multiple surfaces. ToF imaging has been considered a promising technology yet not mature for practical real-world applications as a result of these limitations. However, several contributions have attempted to overcome the disadvantages and apply ToF devices to different purposes: face recognition [MW09], robotics [GPT10, VG10], video surveillance [Sil07, LS10], object detection [RH07], and assisted surgery [OKS15, PPL+16].

3.1.2   Time-of-Flight Image Processing Pipeline

Among all the elements influencing the quality of a range image (dynamic range, contrast, distortion, disparity, motion blur, etc...), three relevant factors are considered to be essential: spatial resolution, intensity resolution, and depth accuracy. Spatial resolution determines the level of detail and the size of the image. Intensity resolution refers to the degree of discernible changes between intensity levels in the image. In turn, the intensity level can be defined as the amount of light in grey scale or the number of colors. Therefore, the larger the intensity resolution the higher amount of brightness or number of colors will be represented in the image. Last, depth accuracy is related to the amount of image noise, as it depends on the precision of distance measurements. Both spatial and intensity resolution are related to the quality of the sensor integrated in the depth scanner, while image noise affecting the accuracy also depends on other factors: object properties (texture, reflectivity, etc.) or scene properties (calibration, ambient illumination, etc.).

The choice of the appropriate 3D acquisition device depends on the application at hand. In the last section, the analysis of the available technology revealed the compromise between cost, accuracy, resolution, and speed. Laser and structured-light scanners have evolved to a point in which they are able to provide high quality 3D representations of objects. Stereo vision systems provide low quality range images at a low cost, yet lack efficiency and most applications require a tedious calibration process. The intention of this first dissertation objective is to explore alternative solutions to complexity in range IR problems while aiming for efficiency. In consequence, ToF devices are considered for providing low quality real-time data at a reasonable cost.

A particular type of ToF sensors are Photonic-Mixer-Devices (PMD) [RMH07], designed by PMD Technologies GmbH and integrated in a consumer camera known as PMD[Vision]® CamCube 2.0. This device features a resolution of a $204x204$ resolution, 25 Hz full-resolution frame rate. The IR illumination unit operates at a wavelength of 870 nm and the receiver integrates a CMOS sensor with suppression of background illumination (SBI) features, which allows to measure distances between 0.3 and 7 meters. The total field of view of the camera optics reaches $40°x40°$. Until PMD Technologies discontinued the CamCube commercial line of products in 2012, this ToF sensor provided the highest resolution in the market.

In terms of quality, range images directly acquired by the ToF camera do not satisfy the minimum requirements to undergo a successful IR procedure. High levels of noise in the images (as depicted in Figure 2) result in severe deformations in the surface of the scanned scene, hindering the alignment of different views of the same object or subject. Under these circumstances, an image preprocessing stage is required to overcome the limitations of ToF range images and enhance their quality. The most frequent tasks within any preprocessing stage are image cleaning and image enhancing, involving denoising (frequency filtering or amplitude thresholding), small component removal, surface smoothing, or motion compensation. Several methods have been specifically designed to enhance ToF range images [PBP08, May09, EOHM10, HLCH12, LKS+13].
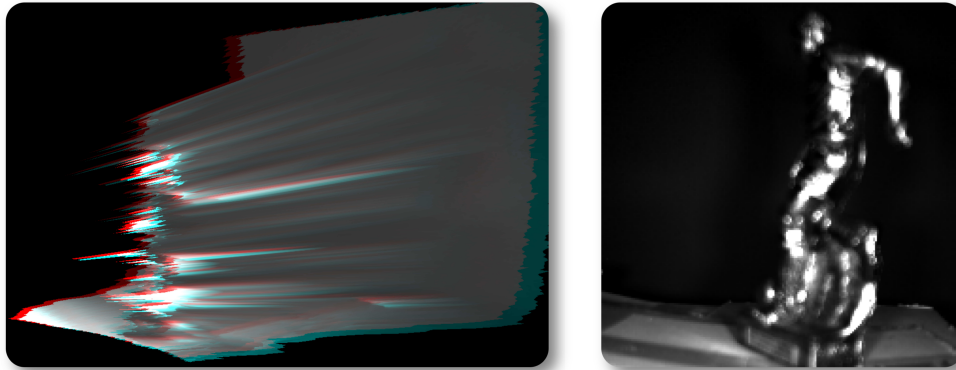

Figure 2: 3D surface representation (depth map) of an object captured using a ToF camera.

A novel approach is proposed within the framework of this thesis with the aim of enhancing ToF image quality. It consists of a three-staged pipeline design reducing the level of noise and extracting relevant information from the images to aid the alignment procedure. A calibration [FH08] is performed during the first stage to correct systematic errors. The ToF camera yields for each frame a depth, an amplitude, and an intensity image. The latter corresponds to the grey scale value of each pixel in the scene, while the amplitude image reflects the amplitude of the signal received by the sensor, which translates to the relative accuracy of the distance measurement. Those distances are automatically discarded by a minimum amplitude filter defined as a statistical threshold in order to remove noisy pixels from the depth image. Once the ambient illumination and systematic amplitude errors are controlled, the first stage comprise the acquisition of an averaged image of the static scene by means of modifying the parameters of the signal. Specifically, the adjusted parameters are: integration time (balancing frame rate and measurement accuracy) and modulated frequency (balancing range/field of view and measurement accuracy). Although the frame rate is heavily reduced, the results of the first stage are noticeable (middle column of Figure 3).

Despite the improvement obtained during the initial stage, distance measurements still derive in surface deformations. Usually, median, amplitude correction, or jump edge filters are applied [FM08, SLP+07, ODCB07] to remove image noise. However, as the range image was already filtered, only a correction for illumination inhomogeneity is applied to prepare the image for the next step. The second stage comprises an enhancement technique based on the shading constraint principle: "If the reflectance properties of the surface are known, a certain range map implies a corresponding intensity image" [BHMB10]. This technique follows the shape from shading (SfS) [ZTC99] which is based on the analysis of the pattern of shades to infer the shape of the surface, exploiting the dependence between range and intensity information. Hence, the shape of a scene is built using local surface normals estimated from the direction of the illumination and the albedo of the surface (which is a measure for reflectance and depends on the surface material). SfS is used as complementary information from the intensity ToF image, correcting noisy depth measurements and enhancing the final quality of the image (right column of Figure 3).
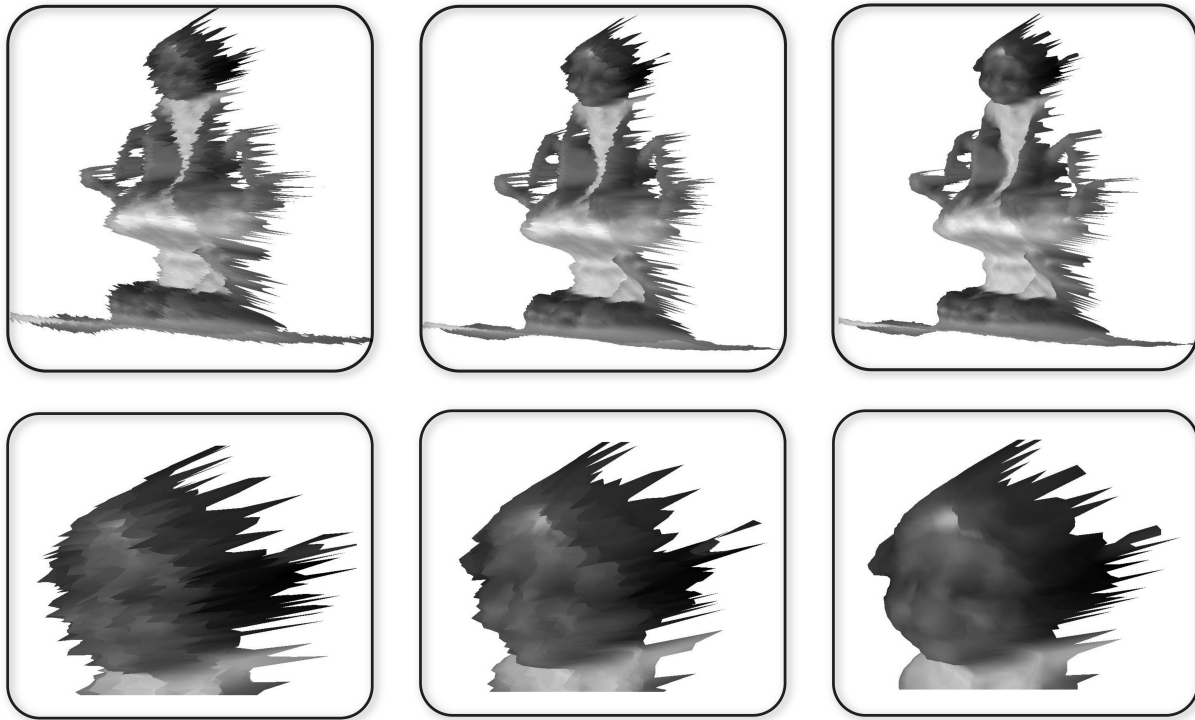
Figure 3: Comparison between the original Tof image (left column), the results of applying the first stage of the pipeline (middle column), and the enhanced image after the second stage (right column).

After concluding the denoising stages, the resulting images are now suitable for the last stage of the pipeline: 3D feature extraction. The third stage aims to extract a reduced subset of informative and characteristic points used during the IR procedure. Across the literature a vast number of feature descriptors have been presented to extract distinctive information from both 2D and 3D images, being the most renowned Harris corner detector [HS88], SIFT [Low04], mesh saliency [LVJ05], and scale dependent corners [NN07], among many others [Kri14]. In order to exploit the information available in both the depth and the intensity images, two subsets of 3D features are considered:

- **Salient points**, a combination of mesh saliency with statistical descriptors proposed by Castellani et al. [CCFM08]. This descriptor extracts a small quantity of disperse interest points with information about curvature, prominence levels, and surface normals. It is a robust feature descriptor able to handle noise, holes, and occlusions.

- **Speeded up robust features** (SURF), one of the most used invariant feature descriptor in CV [BETV08]. SURF is robust to noise, errors, and geometric distortions while handling image blurring and rotation. It is based on an approximation of the determinant of Hessian detector [Lin13]

Once 2D SURF features are detected, the corresponding 3D points are extracted in order to build a homogeneous set of 3D features along with salient points that will guide the feature-based IR method described below. Figure 4 summarizes the preprocessing pipeline.

Figure 4: The proposed ToF image preprocessing pipeline.

3.1.3 Bacterial Foraging for Range Image Registration

Prominent global optimization MHs have been widely studied and successfully applied to tackle IR problems [SPH01, CDB03, MZ0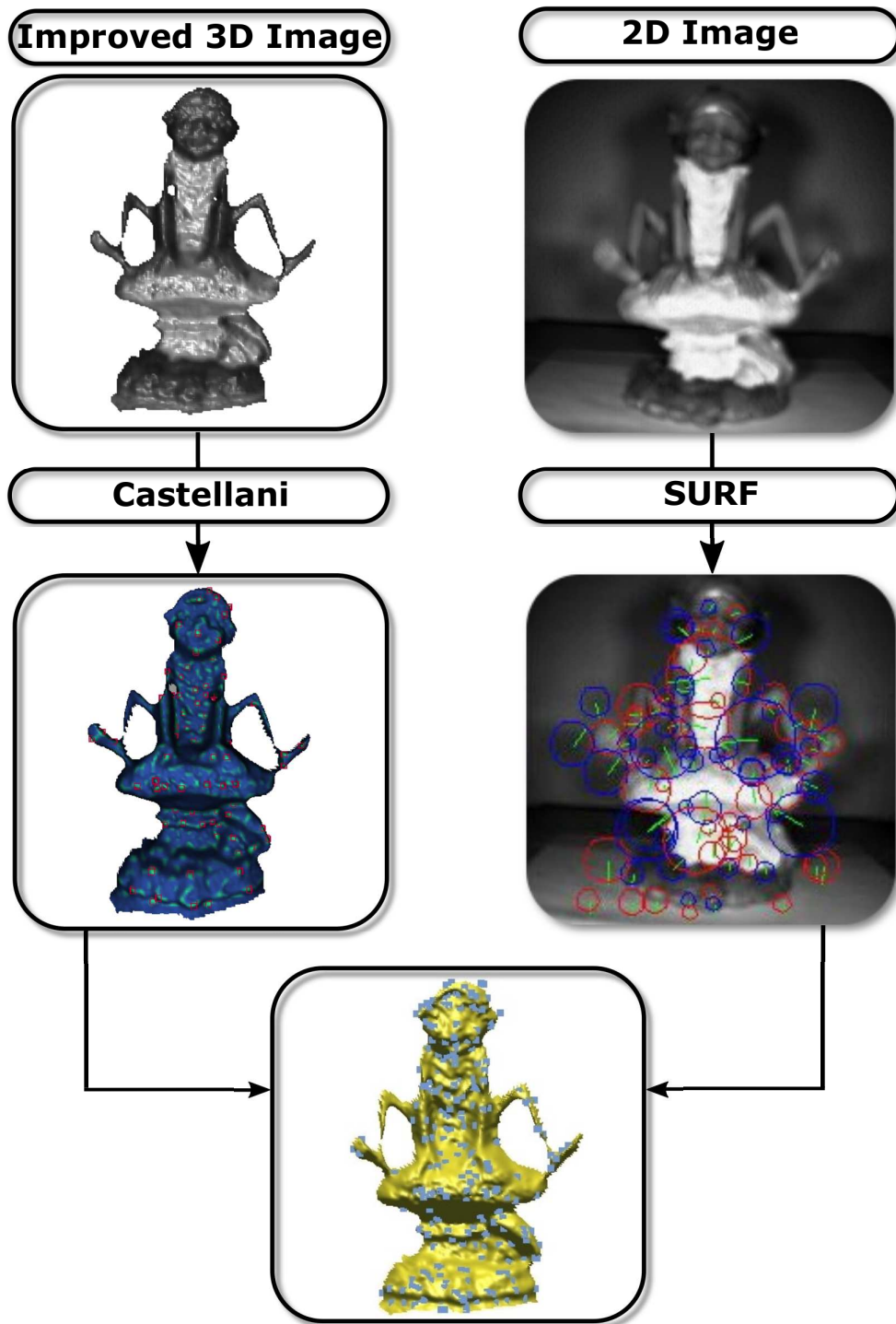4], as well as different bio-inspired approaches, in particular several GAs which are currently considered as part of the state of the art [HN02, CTL04, SBB05]. Recently, another NI proposal has drawn the attention of the optimization community thanks to the good performance and efficiency achieved when solving real-world optimization problems in several application domains [YJH13, DG14, KA15, ZW16, PXZF17]. This approach is known as Bacterial Foraging Optimization Algorithm (BFOA), and it was proposed by Passino et al. in [Pas02]. The source of inspiration of BFOA comes from the social foraging behavior of a bacteria present in the human intestines, Escherichia coli (E. coli).

For many organisms, search and consumption of nutrients are essential activities to ensure their survivability. Group foraging is a key strategy for maximizing the exploitation of nutrients while minimizing exposure to natural hazards or predators. Particularly interesting is the social behavior of E. coli, which is able to perceive the location of nutrient sources based on the gradients of chemicals found in the environment. During foraging, individual bacteria travel around the space by taking small steps while searching for sustenance. E. coli bacteria move during a chemotactic step using their flagella, alternating periods of swims (straightforward locomotion in a determined direction) and tumbles (spin to reorient the direction of the movement). At the same time, bacteria secrete chemicals in the environment used to attract others to the nutrients or repel them to ignore potential risks. Predators and toxic substances may entail a hazard for a bacterium, which can suffer an elimination (death) or a dispersion process (sudden relocation of its position in the environment). Once a bacterium has acquired sufficient energy, it can asexually reproduce and duplicate.

Based on the biological concepts of bacterial foraging, the original optimization algorithm is described as pseudocode in Algorithm 1. The method consists of the following four steps [Pas02]:

- **Chemotaxis:** This process simulates the displacements made by a bacterium during its foraging. Depending on the nutrient concentration, the bacterium swims in a specific direction or it tumbles to change its current direction. Mathematically, $\theta^i(j, k, l)$ represents the ith bacterium in the jth chemotactic, k^{th} reproductive, and l^{th} elimination and dispersal phases. The tumble movement is modeled as in [Pas02]:

$$\phi(i) = \frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}} \tag{I.5}$$

  where Δ indicates a random vector whose direction lies in the range $[-1, 1]$. Thus, the chemotactic step is defined as:

$$\theta^i(j + 1, k, l) = \theta^i(j, k, l) + C(i)\phi(i) \tag{I.6}$$

  where $C(i)$ represents the step size taken during the swim step.

- **Swarming:** A social behavior is simulated in this stage. Those bacteria placed in locations with high amount of nutrients tend to attract other bacteria, while those placed in hazardous zones tend to repel them. This cell to cell attraction-repulsion behavior is modeled as follows:

$$J_{cc}(\theta, P(j, k, l)) = \sum_{i=1}^{S} J_{cc}(\theta, \theta^i(j, k, l)) = \tag{I.7}$$

$$= \sum_{i=1}^{S} [-d_{attract} exp(-w_{attract} \Sigma_{m=1}^{p}(\theta_m - \theta_m^i)^2)]$$

$$+ \sum_{i=1}^{S} [-h_{repell} exp(-w_{repell} \Sigma_{m=1}^{p}(\theta_m - \theta_m^i)^2)]$$

where $J_{cc}(\theta, P(j, k, l))$ is the swarming value to be added to the objective function during the chemotactic step j; S is the total number of bacteria; p is the number of variables (or problem dimension) to be optimized; $d_{attract}, w_{attract}, h_{repell}$, and w_{repell} are different coefficients that represent the attractant and the repellent intensity factors of the bacteria.

Begin BFOA
1   Initialize parameters
2   **For** l=1 **to** N_{ed} **Do** /*Elimination-dispersal loop*/
3   **For** k=1 **to** N_{re} **Do** /*Reproduction loop*/
4   **For** j=1 **to** N_c **Do** /*Chemotaxis loop*/
5   **For** i=1 **to** S **Do** /*Chemotactic step*/
6                    $J(i, j, k, l) \leftarrow J(i, j, k, l) + J_{cc}(\theta^i(j, k, l), P(j, k, l))$ /*Compute fitness*/
7                    $\Delta_m(i), m = 1, ..., p$ /*Generate a random direction*/
8                    $\theta^i(j + 1, k, l) \leftarrow \theta^i(j, k, l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta(i)^T\Delta(i)}}$ /*move bacterium*/
9   evaluate $J(i, j + 1, k, l)$
10   /*Swim*/
11                   $m \leftarrow 0$
12   **While** $m < N_s$ **Do**
13                       $m \leftarrow m + 1$
14   **If** $J(i, j, k, l) < J_{last}$
15   update J_{last}
16                           $\theta^i(j + 1, k, l) \leftarrow \theta^i(j, k, l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta(i)^T\Delta(i)}}$ /*move bacterium*/
17   **Else**
18                           $m \leftarrow N_s$
19   **End-If**
20   **End-While**
21   **End-i-For**
22   **End-j-For** /*End Chemotaxis*/
23   **For** i=1 **to** S **Do** /*Reproduction step*/
24               $J_{health}^i \leftarrow \sum_{j=1}^{N_c+1} J(i, j, k, l)$
25   **End-i-For**
26   Sort bacteria in ascending order of J_{health}
27   Duplicate the best S_r bacteria (lower J_{health}) replacing the worst ones
28   **End-k-For** /*End Reproduction*/
29   **For** i=1 **to** S **Do** /*Elimination-Dispersal step*/
30   **If** $rand() < P_{ed}$
31   Eliminate bacterium and initialize randomly its replacement
32   **End-If**
33   **End-i-For**
34 **End-l-For** /*End Elimination-Dispersal*/
End BFOA

Algorithm 1: Pseudo-code of BFOA.

- **Reproduction:** The least healthy bacteria, i.e. those who find less amount of nutrients during the chemotaxis, die during the reproductive step. On the contrary, the healthiest bacteria are asexually split into two copies. Only the first half of the bacteria is considered for reproduction by replacing the remaining bacteria in order to keep the swarm size constant.

- **Elimination-Dispersal:** The last step consists of the elimination or dispersal of some bacteria based on the simulation of sudden or gradual changes in the local environment, e.g. a rise of the temperature may kill a group of bacteria within a region. This event is simulated by the elimination of some bacteria with a small probability and randomly initializing a replacement.

The main shortcoming of this canonical scheme of BFOA is the oscillation of the bacteria's state when they are close to the optimal values. A considerable number of unnecessary chemotactic steps are usually performed. In order to avoid this, Dasgupta et al. [DDAB09] proposed a refined variant in which the step size is adapted as follows:

$$C = \frac{|J(\theta)|}{|J(\theta)| + \lambda} = \frac{1}{1 + \lambda/|J(\theta)|} \tag{I.8}$$

where λ is a positive constant. Thus, when $J(\theta)$ is large, the step size is accordingly increased ($C \to 1$) and it will be decreased once the bacteria approach the global optima. This adaptive approach outperforms the canonical BFOA as shown in [DDAB09, DBDA09].

3.1.3.1 BFOA variant design and coding scheme for RIR

In order to adapt the BFOA design to the problem at hand, two considerations were made. First, the original proposal suffers different drawbacks, some of them identified in [DDAB09]. In addition, the success of different memetic approaches advises the combination of a global strategy with a local search to exploit the optimization procedure and improve the overall performance of the proposal. Therefore, several alternatives of improvement to be considered within the BFOA scheme are introduced below. All those BFOA variants are built by taking Dasgupta et al.'s proposal [DDAB09] as a base. They have been designed by fixing some specific decisions for some of the components and combining different design choices, described as follows:

- In the original proposal, Δ is a random vector which defines the direction of the tumble step within the range of each bacterium (see Eq. I.5). In order to enhance the social behavior of the bacteria simulated in the original swarming stage and modify the tumble step negative effects, an alternative solution is proposed. With a probability p, a bacterium will follow the progress of another successful bacterium in the population, mimicking its direction.

- It may occur in the original reproduction step that the bacterium with less accumulated life (J_{health}) is not the best bacterium in the swarm. This alternative is based on a probabilistic selection of the best bacteria. Instead of sorting the bacteria in ascending order of their accumulated J_{health} values, a probabilistic filter is applied previous to sorting the bacteria. In first place, the current life value of each bacteria is normalized considering the life accumulated during the chemotaxis. Then, the best fit bacteria to be reproduced are selected with inverse probability of this normalized value. Finally, the bacteria are sorted and the first half reproduce as in the original method. This strategy increases the importance of those bacteria that obtained the best fitness values in the last chemotactic cycles.

- Despite the elimination-dispersal step considers a small probability of application (1/8), the performance of the algorithm usually gets down and it may lose some of the progress achieved in the chemotactic steps. Thus, an alternative using the best bacterium found so far to guide the dispersal step can to minimize the negative impact of killing the best bacterium. Moreover, each transformation parameter has a 30% probability of being similar to the best bacterium.

- A set of elite solutions with the $M = 5$ best solutions found so far is also introduced within the scheme. After each chemotaxis, if the algorithm finds a solution (p_{new}) better than the worst elite solution (e_m), a recombination mechanism is applied considering p_{new} and a random elite solution. Recombination is done by using the BLX$-\alpha$ real-coding crossover operator [ES93]. In case the new solution (r_{new}) is better than p_{new}, the former will replace both the worst elite solution and its parent. Otherwise, p_{new} will replace both e_m and its parent.

- Finally, a local search (LS) strategy is used as a hybridization applied after the BFOA's reproduction step. In particular, two different LS algorithms are considered:

  - *Bound Optimization BY Quadratic Approximation (BOBYQA)* is an iterative algorithm based on the optimization of constrained functions using quadratic approximations [Pow09]. These quadratic models are satisfied by interpolation conditions adjusted automatically by the algorithm. The interpolation points keep constant the separation distance, achieving precise local/global optima.

  - *Crossover Based Local Search* (XLS) [BD01] is based on auto-adaptive crossover operators that fit the parent solutions of the population to generate new trial solutions. Given a parent BFOA solution, XLS randomly selects L bacteria of the population and it generates new ones in the parent neighborhood using crossover operators.

Second, the basic design of the algorithm has to be adapted to solve the IR problem by searching the transformation parameters space. Hence, the coding scheme of the proposed BFOA must specify the representation of a solution and the image primitives used to guide the procedure. In order to align pairs of ToF range images, the proposed IR method aims to find the Euclidean motion that brings the scene image (I_s) into the best possible alignment with the model (I_m). In RIR applications, an Euclidean motion is usually considered based on a 3D rigid transformation (f). This transformation is determined by six or seven real-coded parameters when using either Euler or axis plus angle representation for rotation, respectively [Gos05]. Specifically, Euler-based rotation matrices suffer from the pitfall of the *gimbal-lock* and the specialized literature recommends using the second scheme [SKSK85]. Thus, a rigid transformation is defined as: a rotation $R = (\theta, \ Axis_x, Axis_y, Axis_z)$ and a translation $\vec{t} = (t_x, t_y, t_z)$, with θ and $\vec{Axis}$ being the angle and axis of rotation, respectively. Then, the transformed points of the *scene* are denoted by

$$f(\vec{p}_i) = R(\vec{p}_i) + \vec{t}, \ \ i = \{1, \dots, n\} \tag{I.9}$$

where n is the number of points of the I_s image.

Hence, the RIR procedure can be formulated as an optimization problem developed to search for the Euclidean transformation f^* achieving the best alignment of both $f(I_s)$ and I_m:

$$f^* = \ \arg\min_f \ F(I_s, I_m; f) \ \ s.t.: \ f^*(I_s) \cong I_m \tag{I.10}$$

according to the similarity metric, F, being optimized. The median square error (MedSE) (see Eq. I.3) is usually considered as the F function in 3D model reconstruction [RFL02] due to its robustness in presence of outliers (e.g. acquired noisy range images). Notice that both the F function and the objective function have the same meaning within the optimization process. In order to speed up the computation of the closest point of every $f(\vec{p}_i)$ point, GCP [YAF99] has been considered as indexing structure.

3.1.4 Experiments: ToF Range Image Registration

3.1.4.1 Datasets

The range images used in the experiments were obtained from video sequences acquired using a PMD[Vision]® CamCube 2.0 ToF camera and enhanced by applying the preprocessing pipeline described in Section 3.1.2. Three different objects were considered where each range image corresponds to the scan of objects at 45 degrees of rotation using a turn table. Then, every IR problem instance considers two range images of each object, i.e. facing a pair-wise RIR instance. The considered IR instances comprise a really complex IR scenario due to the strong misalignment and the low overlapping between the two views. Figure 5 shows the snapshots of the three considered objects: *Goblin*, *Teddy*, and *Pirate*. The top row depicts the preprocessed range images.

Within the preprocessing pipeline, a reduced subset of feature range points is extracted using specific extraction methods for 2D [BETV08] and 3D [CCFM08] images. The IR method is run considering these subsets of points in order to achieve a fast convergence. Table I.1 shows the number of points of both the preprocessed range image (raw data) and the extracted features using the third stage of the IP pipeline.

RIR Dataset	*Goblin*		*Teddy*		*Pirate*	
	$0°$	$45°$	$0°$	$45°$	$0°$	$45°$
Raw data	11.468	10.983	13.802	13.641	11.340	10.210
Features	224	228	261	259	254	224

Table I.1:   Number of range points of the raw images and the extracted features for the RIR application.
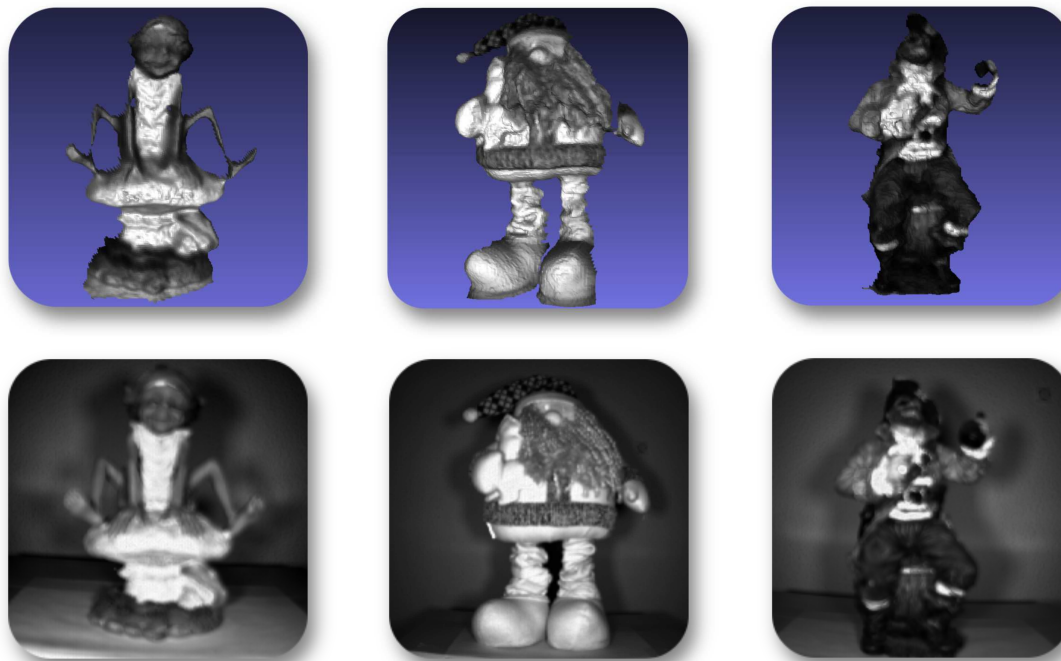
Figure 5:   3D/2D snapshots of the three scanned objects (*Goblin*, *Teddy*, and *Pirate*). Textured range images (left) and their corresponding photo (right) are shown.

3.1.4.2   Experimental Design and Parameter Settings

Table I.2 presents 52 different BFOA variants inspired in the design alternatives introduced in Section 3.1.3.1. Starting from the original proposals, A0 correspond to the Passino's canonical BFOA [Pas02], A1 includes a dynamic evolution of the chemotactic step size to the previous method, and A2 correspond to the Dasgupta et al.'s adaptive BFOA variant [DDAB09]. The only difference between these three methods is calculation of the step size (see Eq. I.8).

Then, four different configuration levels are considered, integrating the previously proposed design alternatives, as follows: first level variants (A3-A6) use Dasgupta et al.'s BFOA as base scheme and integrate one of the proposed alternatives (i.e. the variation in the reproduction step, the alternative direction Δ, the elimination-dispersal step variation, or the introduction of an elite set and the $BLX - \alpha$ operator, respectively). Second level variants (A7-A12) are combinations of two of the proposed alternatives. Then, third level variants (A13-A16) integrate three of the proposed alternatives. The fourth level variant (A17) consists of the combination of the four alternatives proposed in this work.

Finally, the previous set of variants is extended by applying a hybridization to A1-A17 variants using two different LS strategies. Thus, A18-A34 are based on the BOBYQA local search, while A35-A51 make use of XLS. Notice that, when the LS procedure is considered, the BFOA variant becomes a memetic algorithm [Mos89].

Algorithm	Description
Original BFOA methods	
A0	Passino's BFOA [Pas02]
A1	A0 + dynamic evolution of the parameter C
A2	Dasgupta et al.'s BFOA [DDAB09]
First level variants	
A3	A2 + Reproduction variation
A4	A2 + Δ variation
A5	A2 + Elite set and $BLX - \alpha$ operator
A6	A2 + Dispersal variation
Second level variants	
A7	A3 + A4
A8	A3 + A5
A9	A3 + A6
A10	A4 + A5
A11	A4 + A6
A12	A5 + A6
Third level variants	
A13	A3 + A4 + A5
A14	A3 + A4 + A6
A15	A3 + A5 + A6
A16	A4 + A5 + A6
Fourth level variants	
A17	A3 + A4 + A5 + A6
Memetic variants	
A18	A1 + BOBYQA-based LS
A19	A2 + BOBYQA-based LS
A20-A34	A3-A17 + BOBYQA-based LS
A35	A1 + XLS-based LS
A36	A2 + XLS-based LS
A37-A51	A3-A17 + XLS-based LS

Table I.2: Proposed BFOA variants for IR problems.

The performance of the previously described BFOA variants is evaluated in an exhaustive comparison against a high-quality reference set formed by the most relevant evolutionary and MH-based IR methods [DCS11] (See Section 2.2.3):

- Luck et al.'s [LLH00] method (ICP-L).

- Chow et al.'s [CTL04] method (GA-C).

- Silva et al.'s [SBB05] method (GA-S).

- Wachowiak et al.'s [WSZ+04] method (PSO-W).

- Santamaría et al.'s [SCD+07] method (SS-S).

All the IR methods are evaluated considering a two-stage approach as prealignment stage followed by the application of the ICP-based algorithm [Zha94]. To provide a fair comparison, CPU time is considered as stop criterion for the prealignment stage. After a preliminary study, it was determined that 20 seconds is a suitable value once noticed that every method properly converge to accurate enough solutions. In order to avoid execution dependence, 30 different runs of each IR algorithm have been performed in every problem instance. During the refinement stage, the ICP-based algorithm performs 40 iterations to refine the solutions provided by the MH IR methods.

The experiments were performed on an Intel® Core™ i7 2.00 GH platform with 10 GB RAM and implemented using C/C++ and the g++ compiler under GNU/Linux Ubuntu 13.10.

For every IR algorithm, the parameter values recommended by the authors in their contributions were used. Table I.3 shows the parameter values considered in the 52 proposed variants.

Parameter	Value	Description
n_p	7	Problem dimension.
n_s	16	Swarm size.
N_c	30	Number of chemotactic steps.
N_s	20	Number of swims.
N_{re}	6	Number of reproductive steps.
$Prob_{ed}$	0.125	Elimination-dispersal probability.
N_{ed}	3	Number of elimination-dispersal steps.
C_i	0.1	Step length (fixed only for Passino)
λ	0.0001	Constant value used to calculate the step length (C_i) in Dasgupta and HBFOA algorithms.

Table I.3: Parameter settings used in the BFOA-based algorithms.

The experimental analysis is developed from two viewpoints: the numerical optimization of the objective function and the empirical quality of the IR results, considering the prealignment and the refinement stages separately. The combined analysis of results provides more insights regarding to both the optimization and the specific problem-solving branches for each considered dataset.

3.1.4.3   Analysis of Prealignment Results

As seen in the previous section, the considered BFOAs and state-of-the-art evolutionary IR algorithms tackle the prealignment of pairs of range images. We carried out a thorough study of the performance of the different BFOA configurations which is detailed in [BCDS15]. The interested reader is referred to Part II: Section 2, where the complete study is detailed. As a relevant conclusion, memetic configurations demonstrated the best performance of the 52 proposed variants. The use of a local search improved the stability and robustness of the algorithm. Variants integrating XLS obtained better mean performance with respect to those using BOBYQA, which reached the best minimum results.

Next, we analyze the results of using the Kd-tree [Zha94] as closest point rule in the prealignment stage. Table I.4 shows a comparison between the best BFOA variants and the state-of-the-art IR algorithms: ICP-L, GA-C, GA-S, PSO-W, and SS-S. The best performing BFOA methods are memetic: A20, corresponding to a first level variant based in A3 using BOBYQA as LS; A32, a third level variant based on A15 using BOBYQA; A34, a fourth level variant, A17, using BOBYQA; and A49, another variant based on A15 using XLS.

It is appreciable how all the BFOA-based IR methods obtain better mean values than the baseline IR methods. Specifically, A34 and A49 variants obtain the best minimum and mean results in all cases but the *Pirate* instance, in which PSO-W achieves the best minimum outcome. Notice that, this is a significant conclusion showing that the BFOA variants are a suitable choice to design prealignment methods for the 3D model reconstruction pipeline (see Section 2.1). Figure 6 visually compares the registration results of the best IR methods according to the mean performance: A49 and SS-S. In order to present intuitively the visual results, both the model image (blue) and the transformed scene (yellow) have been colored. An accurate registration will result in many colored overlapped regions. Visual results during the prealignment stage show that both A49 and SS-S provide a precise registration for the three considered datasets.

3.1.4.4   Analysis of Refinement Quality

In this section, we extend the previous analysis by means of considering the second refinement stage to provide an accurate solution to the 3D model reconstruction problem. A unified MedSE metric is considered to carry out a fair comparison between both stages. Table I.5 shows the results after the refinement stage, in which it can be observed how the selected BFOA method achieve the best

quality empirical results. All the RIR algorithms provide accurate enough prealignment solutions promoting the proper convergence of the refinement method.

	Goblin			Teddy			Pirate		
	m	μ	σ	m	μ	σ	m	μ	σ
A20	0.00579	0.01076	0.00307	0.01178	0.02367	0.01360	0.00655	0.00770	0.00109
A32	0.00569	0.01075	0.00284	0.01211	0.02034	0.01094	0.00597	0.00742	0.00070
A34	0.00584	0.01089	0.00280	**0.01162**	0.01607	0.00327	0.00606	**0.00725**	0.00041
A49	**0.00561**	**0.00760**	0.00275	0.01226	**0.01413**	0.00140	0.00647	0.00730	0.00052
ICP-L	0.01167	0.01314	0.00027	0.02788	14.7746	20.8366	0.00746	2.74628	0.50859
GA-C	2.49962	303.279	346.837	11.3260	440.990	498.899	3.29864	412.714	520.820
GA-S	0.00947	0.09527	0.07582	0.01638	0.29109	0.38431	0.00598	0.13806	0.37838
PSO-W	0.00842	0.01208	0.00087	0.01710	0.03689	0.01461	**0.00594**	0.00982	0.00276
SS-S	0.00568	0.00971	0.00340	0.01220	0.01446	0.00161	0.00653	0.00738	0.00039

Table I.4: Minimum (m), mean (μ), and standard deviation (σ) MedSE values using Kd-tree in the prealignment stage. Results are scaled ($\times 10^3$).
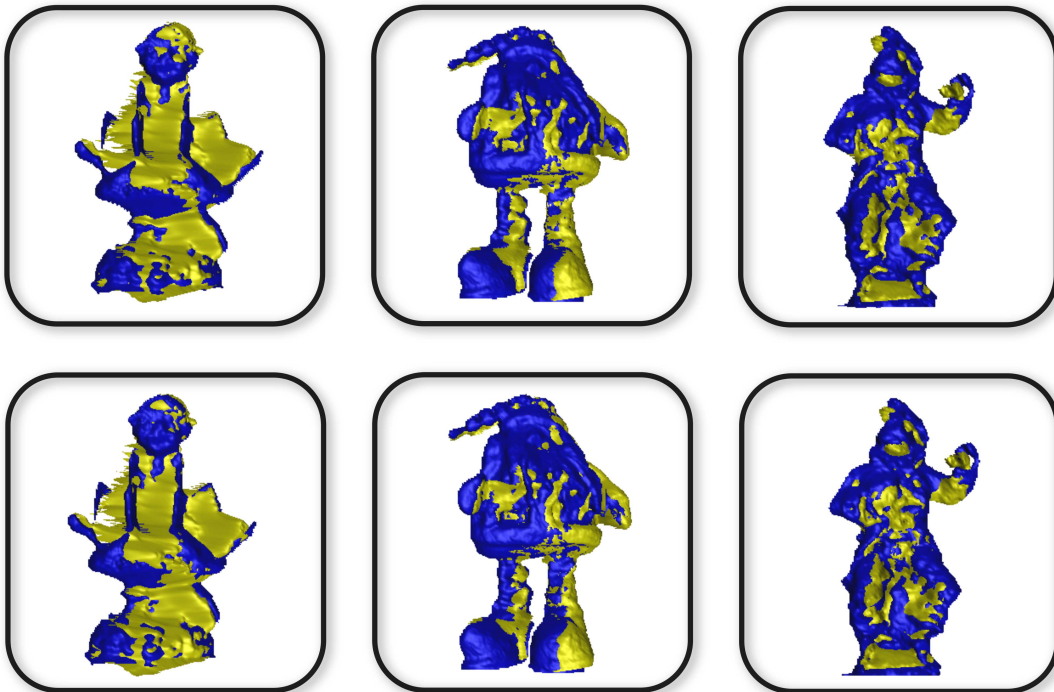

Figure 6: From left to right: First row corresponds to the best RIR estimations achieved by A49 tackling the *Goblin*, *Teddy*, and *Pirate* datasets in the prealignment stage. Second row shows the same results for SS-S.

Specifically, A49 obtains the best mean performance in the three problem instances and the best minimum in the *Goblin* dataset, being the most robust IR method. On the other hand, SS-S matches the minimum result obtained by A49 in the *"Goblin"* dataset, while A22 and A32 achieve the best minimum in the *"Teddy"* and the *"Pirate"* datasets, respectively.

These results demonstrate that the best BFOA variants (e.g. A49) provide a competitive performance compared to the state-of-the-art evolutionary and non-evolutionary IR methods. Figure 7 shows the averaged mean performance of every RIR algorithm in both the prealignment and the refinement stages. The graphic remarks the importance of obtaining an accurate enough prealignment result to promote a suitable convergence of the refinement process. Figure 8 shows the best refinement results for A49. The high-quality reconstruction obtained can be clearly observed.

	Goblin			**Teddy**			**Pirate**		
	m	μ	σ	m	μ	σ	m	μ	σ
A20	0.00520	0.00850	0.00187	**0.01161**	0.01275	0.00167	0.00554	0.00579	0.00010
A32	0.00520	0.00802	0.00201	0.01164	0.01242	0.00132	**0.00545**	0.00578	0.00014
A34	0.00522	0.00844	0.00172	0.01163	0.01202	0.00015	0.00546	0.00575	0.00012
A49	**0.00518**	**0.00611**	0.00179	0.01169	**0.01198**	0.00015	0.00553	**0.00573**	0.00013
ICP-L	0.00934	0.00978	0.00008	0.01215	0.01284	0.00034	0.00579	0.00589	0.00002
GA-C	0.01296	0.07843	0.02832	0.02851	0.46021	0.23071	0.02499	0.07598	0.03782
GA-S	0.00525	0.06017	0.04673	0.01209	0.09637	0.08075	0.00547	0.03242	0.02784
PSO-W	0.00522	0.00942	0.00112	0.01199	0.01295	0.00144	0.00546	0.00586	0.00012
SS-S	**0.00518**	0.00700	0.00180	0.01172	0.01202	0.00015	0.00554	0.00577	0.00011

Table I.5: Minimum (m), mean (μ), and standard deviation (σ) MedSE values using Kd-tree in the refinement stage. Results are scaled ($\times 10^3$).
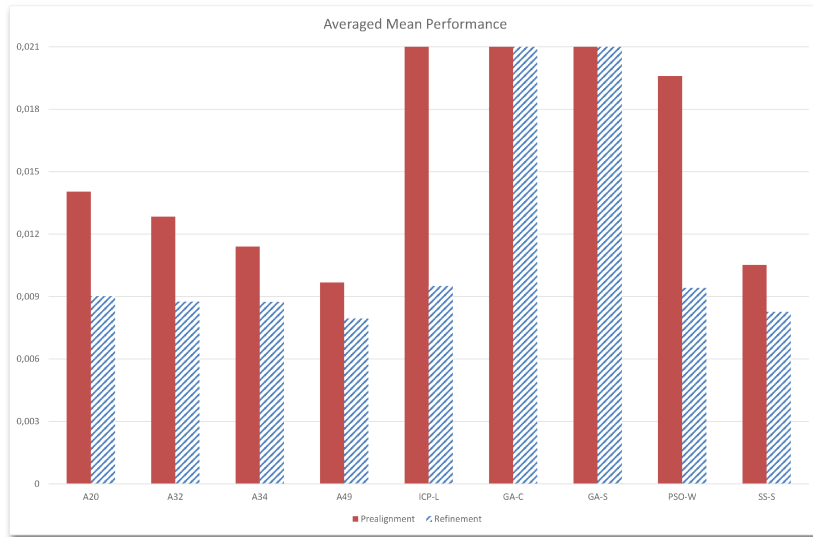

Figure 7: Averaged mean performance of the prealignment and refinement stages obtained in the three considered datasets for both the best BFOA variants and the state-of-the-art evolutionary and metaheuristic IR algorithms. Results are scaled ($\times 10^3$).
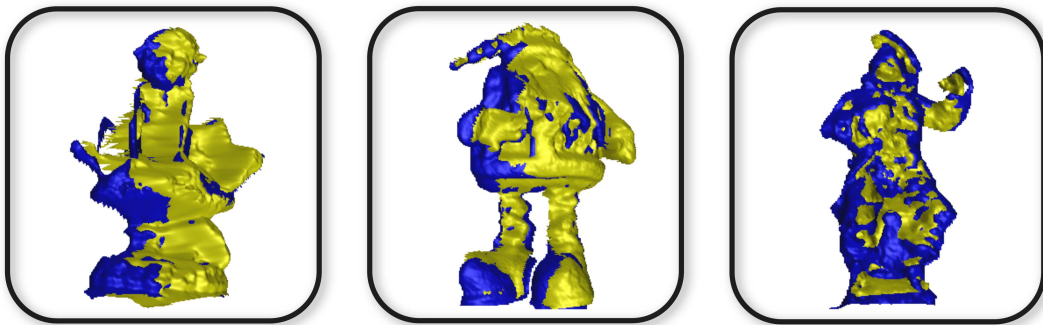

Figure 8: Refinement results of the best RIR estimations achieved using A49.

3.2 Medical Image Registration

Medical imaging (MI) is nowadays a standardized clinical practice. As a result, many specific applications (e.g., assisted surgery) impose both time and accuracy constraints as essential requirements for the techniques involved. IR is a crucial task for any MI application, particularly those involving 3D images. The ultimate goal is to provide IR methods which are in equal terms accurate, reliable, and able to operate in real time. Among the whole spectrum of medical applications, IR techniques involving MI usually can be categorized in one of the following approaches: i) *multi-view registration*, where a scene is scanned from different viewpoints to obtain a 3D representation (e.g., intervention and treatment planning), ii) *multi-temporal registration*, where images are acquired at different times (e.g., disease monitoring), iii) *multi-modal registration*, when images are provided by different sensors or devices (e.g., clinical diagnosis and therapy), and iv) *scene to model registration*, when the scene is compared to a representative model involving different dimension or standard content (e.g., anatomical atlases).

Several contributions have been proposed reviewing the suitability of different medical IR methods (MIR) [HHH01, PMV03, OT14]. In a quest for the precision and reliability demanded by medical applications, the focus is currently centered in the research for intensity-based deformable IR methods able to register multi-modal images. The main disadvantage of intensity-based methods is a highly demanding computational requirement, especially when the approach in question involves the calculation of derivate information (i.e., gradient-based IR methods). As a result, most approaches are not able to operate in real-world constraints due to the complexity generated by considering the intensity information of the images and challenging deformable transformations. In order to overcome the time limitation and improve both the efficiency and the reliability of the procedure, a dimensionality reduction is usually applied in a multi-stage procedure. Several methods rely on the initial application of a rigid transformation before the deformable registration is applied [ZHD$^+$12, CKA$^+$12, LCH$^+$13, LSLO14]. Nevertheless, despite continuous advances in computing power, most approaches still require a considerable time to perform the registration, e.g., around 4 minutes (involving rigid and deformable registration) for the reconstruction of a 3D model of the aorta in surgery planning applications [LSLO14]. Not long ago, a complete global registration procedure of two 3D medical images could take around 30 minutes [JS01].

3.2.1 Bacterial Foraging for Medical Image Registration

As part of this PhD dissertation, the second objective involves the evaluation of the proposed BFOA-based IR method within a complex MI application. Considering the time limitations of intensity-based methods, our proposal follows the feature-based approach as a way of reducing the computational requirements of the problem at hand. Therefore, the proposed IR method fits with the previous trend of considering a fast and lightweight rigid prealignment prior to the application of an accurate deformable MIR approach. In particular, a multi-temporal registration approach is followed, tackling a complex monomodal problem involving different levels of noise and lesions.

Similar to the RIR procedure, feature-based MIR approaches rely on a feature extraction procedure prior to the alignment of the images. Due to the different acquisition techniques employed, 3D medical images contain a lower level of detail than range images, hindering the process of extracting relevant feature points. Common feature descriptors such as Harris corner detector or scale-dependent corners provide insufficient or misleading information. Therefore, feature extraction methods able to detect anatomical information are usually considered [CWS$^+$03, LSD04, QH07].

In particular, the problem addressed in this application involves the registration of brain MRI scans. Sulci or cortical surface information are commonly used for brain registration. To exploit meaningful anatomical information while maintaining efficiency, a crest-line edge detector

[YBS05] is considered. This feature descriptor extracts the isosurface from points with relevant curvature information [MBF92] procuring a set of ridge features (i.e., lines of extreme curvature along surfaces).

The performance of all the 52 BFOA variants proposed for RIR is also evaluated for this MIR application. The internal design of these variants remains practically unchanged. The only modification is related to the considered transformation model. Hence, MIR problem instances consider a 3D similarity transformation (f'), adding one more real-coded parameter to the representation, the scaling factor [Gos05]. Thus, f' will be defined as a rotation R, a translation $\vec{t}$, and a uniform scaling s, involving 8 real-coded parameters to be estimated ($x = \langle \theta, Axis_x, Axis_y, Axis_z, t_x, t_y, t_z, s \rangle$). Then, transformed points of the *scene* are denoted by

$$f'(\vec{p_i}) = sR(\vec{p_i}) + \vec{t}, \quad i = \{1, \ldots, n\} \tag{I.11}$$

3.2.2 Experiments: 3D Medical Image Registration

3.2.2.1 Datasets

The medical images used to evaluate the performance of our proposal (see Figure 9) were extracted from the BrainWeb public repository [1] from the McConnell Brain Imaging Centre [KEP99]. The BrainWeb repository provides synthetic MRI data computationally generated, based on two anatomical models: normal and multiple sclerosis (MS). Such synthetic MRIs have been extensively used by the neuroimaging community to evaluate the performance of different methods [WSZ+04, BHB01, CMG+10, SLP06].

(a)   (b)   (c)
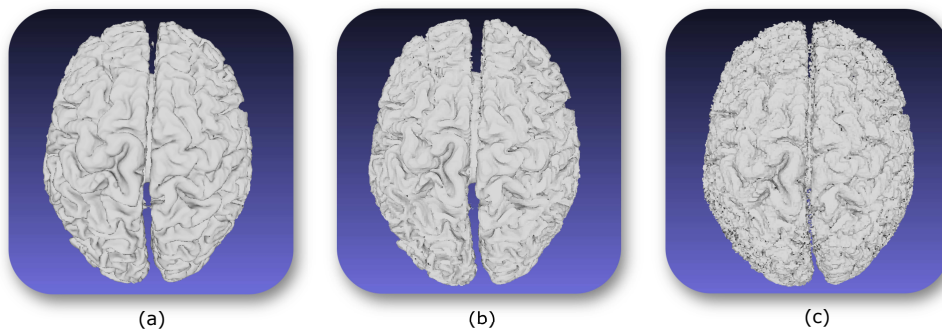
Figure 9: MRIs extracted from the Brainweb image dataset. Noise levels correspond to 0% (a), 1% (b), and 5% (c) noise.

Table I.6 describes the particular settings considered to generate every medical image in our study (named MRIa, MRIb, and MRIc), detailing the nature and the crest line features extracted from each data set. In this application, each medical IR problem instance considers two of the MRIs.

Medical IR Dataset	MRI model	Anatomical modality	level of noise	Raw data	Crest-line Features
MRIa	Normal	T1	0	12.644	584
MRIb	Normal	T1	1	11.493	393
MRIc	Lesion	T1	5	10.112	284

Table I.6: Detailed description of the Brainweb image dataset and the extracted number of features.

---

[1] Available at http://www.bic.mni.mcgill.ca

3.2.2.2   Experimental Design and Parameter Settings

Each medical IR problem instance to be tackled consists of the registration of two medical images (MRIa-MRIb, MRIa-MRIc, and MRIb-MRIc), having three different scenarios overall. For this MI application, we analyze the behavior of the 52 BFOA variants (Table I.2) compared to the most relevant state-of-the-art evolutionary and metaheuristic medical IR algorithms [DCS11]:

- Luck et al.'s [LLH00] method (ICP-L).

- He and Narayana's [HN02] method (GA-HE).

- De Falco et al.'s [DDMT08] method (DE-EA).

- Wachowiak et al.'s [WSZ$^+$04] method (PSO-W).

- Santamaría et al.'s [SCD$^+$07] method (SS-S).

As for the previous experimental design, we considered the parameter settings originally proposed by the authors in every contribution and the same criteria shown in Section 3.1.4.2. In this case, the refinement stage is not considered as it is not usually applied for MIR applications [PMV03]. To design a challenging scenario, a random (similarity) transformation is applied to the ground-truth image for each individual run of the algorithms and then the IR method estimates the unknown inverse transformation. Thus, a different ground-truth registration is available for each MIR problem instance in each of the 30 runs performed by all the IR methods. In particular, similarity transformations are generated following a uniform probability distribution where each of the three rotation axis parameters will be in the interval $[-1, 1]$; the rotation angle will range in $[0°, 360°]$; the three translation parameters in $[-40mm, 40mm]$; and the uniform scaling in $[0.5, 2.0]$.

3.2.2.3   Analysis of Results

In this section we present the summarized results for the medical IR application. The analysis is focused on both numerical optimization and empirical quality viewpoints. The performance of all the 52 BFOA variants proposed is analyzed and detailed in [BCDS15]. The interested reader is referred to Part II: Section 2, where the complete study is detailed. The conclusions of this preliminary study stress the importance of the memetic configurations including a local search, which improved considerably the performance of the original design.

Table I.7 shows a comparison between the best BFOA variants and the state-of-the-art MIR algorithms: GA-HE, DE-EA, PSO-W, and SS-S. The best performing BFOA variants are all memetic: A21, corresponding to a first level variant based in A4 using BOBYQA as LS; A34, a fourth level variant, A17, using BOBYQA; A38, a first level variant based in A4 using XLS; A40, a first level variant based in A6 using XLS; and A46, a second level variant, A12, using XLS.

	MRIa-MRIb			MRIa-MRIc			MRIb-MRIc		
	m	μ	σ	m	μ	σ	m	μ	σ
A21	0.35060	0.74191	0.27962	0.52307	0.86755	0.17151	0.59876	0.88588	0.14071
A34	0.35269	0.76000	0.27096	0.53355	0.87443	0.16380	0.59195	0.86972	0.14687
A38	0.35583	0.70046	0.24895	0.55031	0.79885	0.16414	0.61317	0.89790	0.11956
A40	0.42245	0.72494	0.24390	0.61963	0.83566	0.14476	0.66457	0.82989	0.12712
A46	0.35793	0.61639	0.23701	0.54491	0.82812	0.15644	0.63730	0.86563	0.12452
GA-HE	0.67370	0.92009	0.07014	0.71123	0.93042	0.05905	0.81940	0.94485	0.03678
DE-EA	0.35038	**0.45045**	0.21427	0.52131	**0.68169**	0.19055	0.57298	**0.74460**	0.17087
PSO-W	0.98314	0.99540	0.00410	0.98765	0.99617	0.00359	0.99127	0.99729	0.00222
SS-S	**0.34941**	0.52681	0.27068	**0.52129**	0.70433	0.20847	**0.56538**	0.76076	0.18465

Table I.7:   Minimum (m), mean (μ), and standard deviation (σ) MedSE results of the best BFOA variants and the state-of-the-art evolutionary IR algorithms, considering the Brainweb MRI dataset.
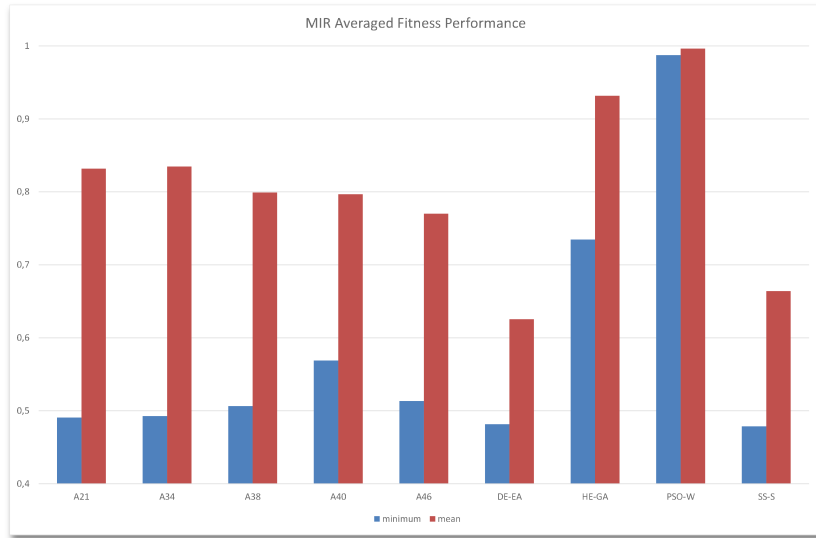
Figure 10: Averaged minimum and mean MedSE results of the three Brainweb MRI instances for both the best BFOA variants (A21, A34, A38, A40, A46) and the state-of-the-art evolutionary and metaheuristic IR algorithms.

The best BFOA variants obtain competitive results compared to the best state-of-the art IR methods, DE-EA in average and SS-S in minimum results. Figure 10 depicts the averaged performance of the three instances regarding to the minimum and the mean results, respectively.

Next, we analyze the final registration quality between the original images considering the unified MeanSE metric, more suitable for MIR applications. This metric allows us to include ICP-L in the comparison, as it is not driven by the similarity metric used in the previous analysis. Table I.8 shows the quality results of the best BFOAs compared to the state-of-the-art evolutionary and MH-based algorithms. In this case, BFOA variants A34 and A38 obtain the best minimum results in two instances, and DE-EA in the other one. Comparing the average performance, the results show how the best BFOA variants are usually the third best IR method, achieving competitive results with the most robust IR method, DE-EA, and outperforming ICP-L, GA-HE, and PSO-W. Figure 11 visually compares the results for the most robust BFOA variant, A46, and the best state-of-the-art IR algorithm, DE-EA.

	MRIa-MRIb			*MRIa-MRIc*			*MRIb-MRIc*		
	m	μ	σ	m	μ	σ	m	μ	σ
A21	7.52734	33.6204	30.5569	15.8257	40.9494	31.0017	14.5219	35.2496	29.6970
A34	7.54181	31.8119	31.6494	15.6649	40.7653	36.4613	**14.3074**	33.8253	30.6842
A38	7.58665	19.9195	19.5529	**15.6354**	27.1802	22.8515	14.7088	33.8536	24.3227
A40	7.73652	26.6422	36.9432	15.7340	34.8252	35.8491	14.3778	24.0626	20.5123
A46	7.58779	17.7286	19.8932	15.9884	27.4294	21.2792	14.3603	26.8315	20.2831
ICP-L	7.53203	75.2935	96.9296	15.7313	77.1457	64.7272	15.4881	38.5663	24.7864
GA-HE	10.9343	42.2374	53.9668	15.9136	37.8208	28.4723	14.7316	33.4822	38.7102
DE-EA	**7.52046**	**10.0073**	5.52872	15.8603	**19.0766**	5.06875	14.7249	**18.7715**	5.94312
PSO-W	115.685	2283.250	2215.23	70.3427	2353.43	2183.17	77.1475	2183.79	2130.71
SS-S	7.52120	14.64251	15.3293	15.9470	26.0639	21.3367	14.8053	20.9225	9.75530

Table I.8: Minimum (m), mean (μ), and standard deviation (σ) MedSE values using Kd-tree considering the Brainweb MRI dataset.
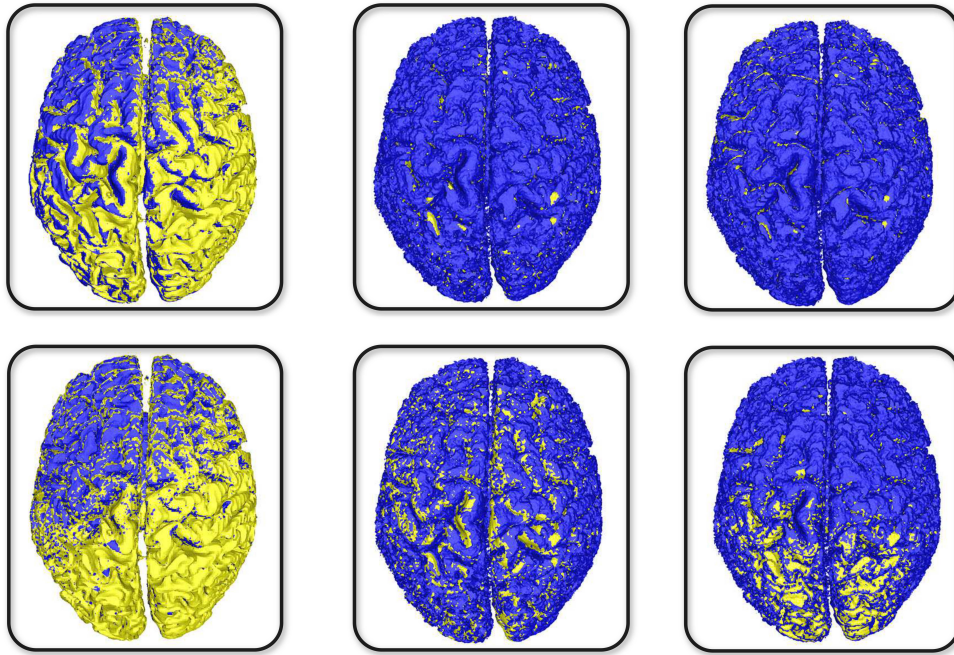
Figure 11: From left to right: First row corresponds to the best medical estimations achieved by A46 tackling the *MRIa-MRIb*, *MRIa-MRIc*, and *MRIb-MRIc* instances, respectively. Second row shows the same results for DE-EA.

These results demonstrate that the best BFOA variants (e.g. A46) provide a competitive and robust performance with respect to the state-of-the-art evolutionary and MH-based MIR methods, although at a significant distance of SS-S and DE-EA. Figure 12 shows the averaged mean performance of every algorithm over the three MIR scenarios. Notice that all the BFOA variants outperform ICP-L, GA-HE, and PSO-W in every case, with A46 being the best performing one.
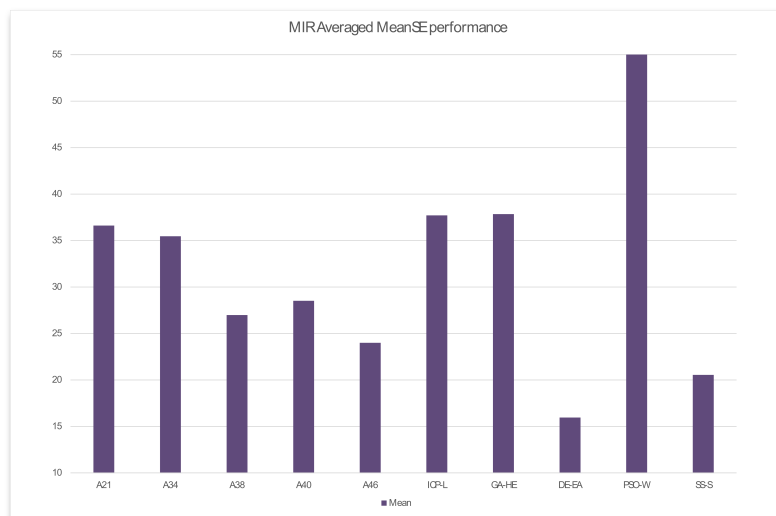

Figure 12: Averaged MeanSE results of the three Brainweb MRI instances for the best BFOA (A21, A34, A38, A40, A46) and the state-of-the-art evolutionary and MH-based IR algorithms.

3.3 Skull-Face Overlay and Articulated Image Registration

Craniofacial superimposition (CFS) [Yos12] established its foundations in the anthropological studies of Broca and Bertillon [Bro75, Ber96] by the end of the 19th century, who studied the relationship between cranial structures with the corresponding facial features and introduced the collection of a suspect's physiognomy for criminal records (forensic photograph), respectively. Ever since, CFS has matured along with digital imaging developments setting significant milestones for craniofacial and forensic identification.

The earliest documented use of CFS in criminal cases can be traced to the beginning of the 20th century [Ube99]. One of the most relevant examples is known as the Ruxton case (1937), where the skulls of two missing women in Scotland were compared with their ante-mortem (AM) photographs adjusted to life size. The process involved the superimposition of the image tracings (negatives) on an X-ray scan of the skull and a subsequent comparison to corroborate the correspondence before court. The procedure was labelled as photographic superimposition [GB37]. During the following 50 years, the basis of modern forensic studies was revised with different anthropological measurements, indices, and features [MS56, Maa89, Geo93], while the use of CFS spread as identification technique under adverse conditions [Sen62, BH89, Ube00, Ind09].

The introduction of video cameras in the superimposition process involved a major achievement for CFS. Video superimposition greatly simplified the process by reducing the time and effort required to adjust the orientation of the skull to fit the pose in the photograph. Several proposals followed the application of this novel approach since the late 1970s [HG77, Koe82, BDW86, SY93, JSA01, FHS08], which burst into popularity after the identification of Josef Mengele (Hitler's chief medical officer) [Hel86].

1991 was another turning point for CFS with the appearance of the first "automatic" computer-aided method. Until then, most computer-assisted approaches were, in essence, photographic or video superimposition methods where computers were used as data storage or simple visualization units [LC85, DCV$^+$86, UBO92]. Nickerson et al.'s proposal [NFKC91] involved the digital modeling of the surface of a skull and a facial photograph. The method was based in a GA that estimates a near-optimum superimposition between the 3D model and the photograph, thus reducing the problems of previous approaches regarding scaling and a correct estimation of the overlapping. Many other approaches have been proposed exploiting the capabilities of digital devices and computer graphics [GS01, BSA05, BCD$^+$07, SCD$^+$07].

In an attempt to standardize the CFS technique, recent studies [DCI$^+$11, DWK$^+$15, DC18] categorize the diversified procedures followed to perform the superimposition, examine its drawbacks and limitations, and establish a series of guidelines and best practices. Among the existing methods, three main categories are distinguished: photographic, video, and computer-aided methods [HIWK15]. Meanwhile, computer-based methods are in turn divided in photo, video, and 3D/2D computer-aided CFS methods.

Three consecutive stages for the whole CFS process have been distinguished in [DCI$^+$11]: 1) The acquisition and processing of the materials, i.e., the AM facial images and the skull (or 3D model of the skull). In some approaches, this step also involves the location of the craniometric and facial landmarks on the skull and the face, respectively. Diverse IP techniques can also be applied during this stage to enhance the quality of the images [GW08]. 2) The skull-face overlay (SFO), which focuses on achieving the best possible superimposition of the skull and a single AM image of the missing person. This process is iteratively repeated for each available AM image, obtaining different overlays. Skull-face overlay thus refers to what traditionally has been known as the adjustment of the skull size and its orientation with respect to the facial photograph [Yos12, Jay15]. It is the most time-consuming stage of the whole CFS procedure. 3) The decision making, where a degree of reliability for the craniofacial correspondence is established in order to show whether the skull and the available photograph belong to the same person or not (exclusion). This task

requires a thorough analysis of the face/skull correspondence provided by SFO to determine if the skull and the face actually belong to the same person [DWK$^+$15].

Regardless the technical category followed by a particular approach, most methods rely in the comparison of anatomical landmarks located in the skull and the face. These landmarks are located in two different objects (a skull and a photograph) resulting in variable correspondences due to the thickness of soft tissue which depends on age, weight, and facial expression. While different contributions have been proposed suggesting different sets of landmark configurations or locations, Martin and Saller [MS56] proposal is frequently considered (see Figure 13).
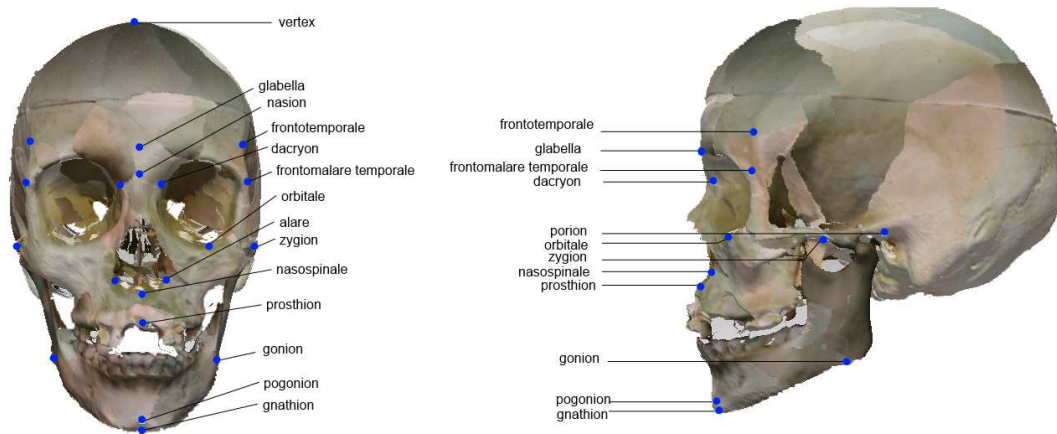

Figure 13: Cranial landmarks in frontal and lateral view.

There is a strong interest in designing automatic methods to support the forensic anthropologist to put CFS into effect. In particular, the design of computer-aided CFS methods has experienced a boom over the past twenty years [HIWK15]. The most recent approaches use 3D models of the skull and SC methods for the first two CFS stages [IBC$^+$09, ICDS11, CÁInCA$^+$15, CÁCÁG$^+$17] and include CV techniques to support the final decision [CÁIC16, CÁICW18]. Next section is devoted to review the evolution of computer-aided SFO methods belonging to the second stage and considering 3D/2D superimposition. These methods allow the automation of different tasks and handle the inherent uncertainty of the CFS procedure. In particular, the appropriate projection of the skull over the facial photograph, which is a very challenging and time-consuming part of the CFS technique. In fact, the estimation of the best correspondence for the SFO can take hours to reach the best possible fit [FHS08, Ube00].

Up to now, all computer-based SFO methods have considered the mandible as a rigid (i.e., not articulated) part of the skull. Such a simplification causes a negative impact on the accuracy of the automatic SFO method. As the AM images used to perform CFS are typically provided by relatives, the missing person usually appears in relaxed situations, most of them smiling or with the mouth slightly open. Generally, cases with grimaces or forced poses are discarded due to the fact that the mandible is in an exaggerated position and these kinds of facial expressions distort the soft tissue of the face. Moreover, those photographs where subjects appear with their mouths open reduce the confidence of the identification. Therefore, it is essential to model the articulation of the mandible in order to improve CFS reliability, considering that in most cases we only have skeletal information available to infer its movement. Moreover, the need to reproduce the position of the mandible as displayed in the AM photograph has been identified within the best practices in the CFS technical procedures [DWK$^+$15] produced by the MEPROCS international group[2]. Usually,

---

[2]The FP7 European project entitled 'New Methodologies and Protocols of Forensic Identification by Craniofacial Superimposition' (MEPROCS) aimed to develop a common methodology for the application of CFS, facilitating the application of this technique in forensic practice.

a forensic expert proceeds to locate the mandible in order to approximate the facial expression of each one of the available photographs, either manually or using 3D modeling software.

3.3.1   Skull-Face Overlay: an Image Registration Problem

From a general point of view, the SFO problem involves positioning the skull in the same pose as the face in the photograph. In terms of CV, the AM photograph is the result of the 2D projection of a real (3D) scene that was acquired by a particular (unknown) camera [FV93]. In such a scene, the living person was in a given pose somewhere inside the camera's field of view. The most natural way to deal with the SFO problem is to replicate that original scenario. To do so, a 3D model of the skull must be used. The goal of this process is to adjust the size and the orientation of the skull model with respect to the head in the photograph [CÁIN+14b]. Besides, the specific characteristics of the camera must also be replicated to reproduce the original situation as much as possible. The latter is the approach followed by state-of-the-art automatic methods. In these approaches [IBC+09, ICDS11, CÁInCA+15], the 3D skull model is positioned in the camera coordinate system through geometric transformations, i.e. translation, rotation and scaling, which corresponds to the adjustment of the skull size and its orientation at the same angle as the face in the image [Yos12]. Finally, a perspective projection of the 3D skull model is performed onto the facial photograph. This is modeled as a 3D/2D image registration (IR) problem [IBC+09] whose solution is guided by a set of cranial and facial landmarks previously located by a forensic expert on both the 3D skull model and the facial photograph (see Fig. 14).

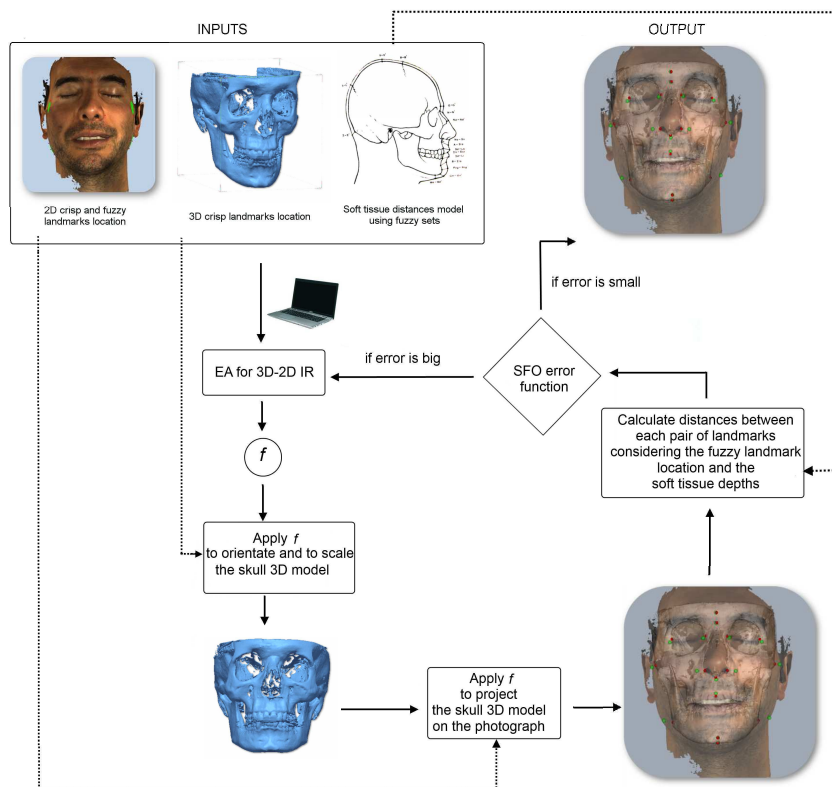Figure 14: Scheme of the SFO procedure as a 3D/2D IR problem [CÁIN+14b]

The 3D/2D IR task involved in SFO aims to reproduce the original scenario of the facial photograph while acknowledging two sources of uncertainty [IBC+09]: the camera configuration at the moment of the acquisition (focal length, distance from the camera to the individual), and the pose of the skull 3D model (orientation, resolution, and size).

Once the location of these landmarks is provided by the forensic anthropologist, the SFO procedure is focused on searching for the skull orientation leading to the best matching of the two sets of landmarks. The required perspective transformation that needs to be applied to the skull was modelled in [IBC$^+$09] as a set of geometric operations involving 12 parameters/unknowns encoded in a real-coded vector to represent a superimposition solution.

Hence, given two sets of cranial and facial landmarks, $C = \{cl^1, ..., cl^n\}$ and $F = \{fl^1, ..., fl^n\}$, the overlay procedure aims to solve a system of equations with the following 12 unknowns: the direction of the rotation axis $\vec{d} = (d_x, d_y, d_z)$, the location of the rotation axis with respect to the center of coordinates $\vec{r} = (r_x, r_y, r_z)$, the rotation angle θ, the factor s that scales the 3D skull model as the face in the photograph, the translation $\vec{t} = (t_x, t_y, t_z)$ that places the origin of the 3D skull model in front of the camera to replicate the same pose of the photograph, and the camera's field of view ϕ.

The authors employed different EAs to address the optimization process. Additionally, they extended the original objective function which minimizes the distances among cranial and facial landmarks to accommodate landmark location uncertainty (resulting from the difficulty to locate the facial landmarks in their exact locations using a photograph) [ICDS11] and landmark matching uncertainty (corresponding to the soft tissue distances existing between facial and cranial landmarks) [CÁInCA$^+$15] using fuzzy sets. As a result, the objective function to be minimized is defined as the Fuzzy Mean Error (FME):

$$FME = \frac{\sum_{i=1}^{Ncrisp} (d'(x_i, f(\tilde{C}^i)) + \sum_{j=1}^{Nfuzzy} (d''(\tilde{F}^j, f(\tilde{C}^j))}{N}, \tag{I.12}$$

where *Ncrisp* is the number of 2D facial landmarks precisely located (crisp points), *Nfuzzy* is the number of 2D facial landmarks imprecisely located and defined as 2D fuzzy sets, N is the total number of landmarks considered ($N = Ncrisp + Nfuzzy$), x_i corresponds to a 2D facial landmark defined as a crisp point ($x_i \in F$), $\tilde{C}^i$ and $\tilde{C}^j$ are fuzzy sets modeling each 3D cranial landmark and the soft tissue distance to the corresponding 3D facial landmark i or j; f is the function that determines the 3D/2D perspective transformation that properly projects every 3D skull point onto the 2D photograph; $f(\tilde{C}^i)$ and $f(\tilde{C}^j)$ are two fuzzy sets, corresponding to the result of applying the perspective transformation f to the 3D volume, which models the landmark matching uncertainty; $\tilde{F}^j$ represents the fuzzy set of points of the imprecise 2D facial landmark; $d'(x_i, f(\tilde{C}^i))$ is the distance between a point and a fuzzy set of points, and $d''(\tilde{F}^j, f(\tilde{C}^j))$ is the distance between two fuzzy sets. Note that once estimated, f is applied to every 3D cranial landmark, handling the skull as a single rigid (i.e., not articulated) object.

3.3.2 Automatic Estimation of Mandible Aperture for Skull-face Overlay

In order to adequately model the mandible articulation in automatic SFO, it is essential to thoroughly analyze the mechanics of its movements. In Section 3.3.2.1, we review the existing mandible aperture models. Given the conditions to perform SFO, we select the best aperture model for our aim in Section 3.3.2.2. Hence, this section is conceived as introduction of the prerequisites involved in the design of the articulated SFO method proposed afterwards in Section 3.3.3.

3.3.2.1 Mandible aperture modeling
The Temporomandibular Joint (TMJ) connects the mandibular condyle to the temporal bone in the cranium, which are separated by a small oval fibrocartilage called articular disc [Wil11]. This articulation is one of the most complex joints in the human body [Hyl92] as it features a bilateral articulation that can perform both gliding (translation) and hinge (rotation) movements [Mag14, MGL$^+$09]. Both the articulation muscles and the articular disc allow the mandible to

perform the different movements, while teeth occlusion, ligaments, and bones constrain the degrees of freedom of the joint [AMC$^+$07] (see Fig. 15).
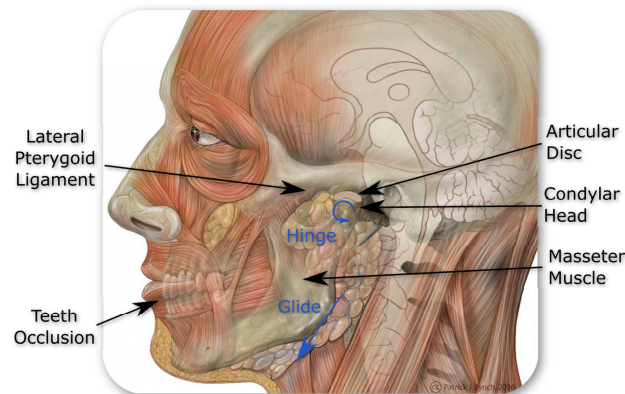

Figure 15: Illustration of the anatomy of the TMJ, by Patrick J. Lynch, medical illustrator, C. Carl Jaffe, MD, cardiologist. Used under Creative Commons Attribution 2.5 license 2006.

This complex movement has been widely studied in the oral biology and maxillofacial surgery fields. Most of the existing approaches try to represent the movement using biomechanical models [GLK$^+$95, XBP$^+$08, OVBG97, KV97, KvE95], involving the stress forces induced by the muscles and the ligaments of the joint. Others attempt to explain this movement using a mathematical kinematic model, e.g. defining a kinematic axis [KUS93], a kinematic center [YZMN95], an instantaneous center of rotation (ICR) [Che98], a screw axis [GGM$^+$00], a helical axis [GAAP97], or a non-orthogonal floating axis [LBDR03]. Still, none of the latter proposals is able to faithfully simulate the anatomical joint, as all of them have limitations related to the geometric constructions and their calculations or inconsistences when determining the rotational axis, dealing with wide variations in the reproducibility of the opening movements [YZMN95, CK99].

In Lemoine et al. [LXA$^+$07] the authors used a 3D optical tracker to record mandibular movements during mouth opening and closing, and studied the relations of the mandibular geometry to define a prediction algorithm that can be applied to different case studies. They found a linear relationship between the length of the mandibular radius (i.e., the line between the center of the condylar head and the lower central incisor) and the mandibular movement. Thus, the mandibular radius is used as the scaling factor to fit the prediction model to each particular individual. Even though their experiments are based on a small sample of individuals, the results of their model concur with other studies based on recording in-vivo measurements [BGE$^+$13] or the proposal of a hierarchical ICR model to predict the articulation movement [VNFM12].

3.3.2.2 Selected mandible articulation model

In choosing the model, the main concern was the applicability to our SFO scenario. Even at the expense of accuracy loss, the selected mandible articulation model cannot be based on the interaction of the soft tissue, i.e. the position of the ligaments or the strength of the facial muscles. On the contrary, the movement of the mandible in SFO must be modeled using information that can be extracted from the bones only.

Lemoine et al.'s approach [LXA$^+$07] offers a good compromise between reliability and simplicity when considering the parameterization of the model. In particular, mandible movement is modeled as a composition of: i) a rotation around the condylar axis (defined between the center of the two condylar heads) followed by ii) a translation along the antero-posterior direction and iii) a supero-inferior translation. The magnitude of the three movements depends on the aperture

percentage (with 0% being teeth occlusion and 100% being maximum opening according to the normalized path shown in Fig. 16). It is important to highlight that up to 75% of the maximum opening may be approximated considering just the aforementioned rotation.

The magnitudes of the different movements are scaled using the mandibular radius, which can be easily identified in a radiographic analysis, measuring the distance from the center of the condylar head (condylar center) to the lower central incisor. In order to scale the normalized path, two linear regression equations are defined within the prediction model, relating the mandibular radius R_M with the maximum mandibular rotational angle (Θ_M) :

$$\Theta_M = (2.42)R_M - 1.24 \tag{I.13}$$

and with the maximum mandibular translation (T_M):

$$T_M = (-0.165)R_M + 2.91 \tag{I.14}$$
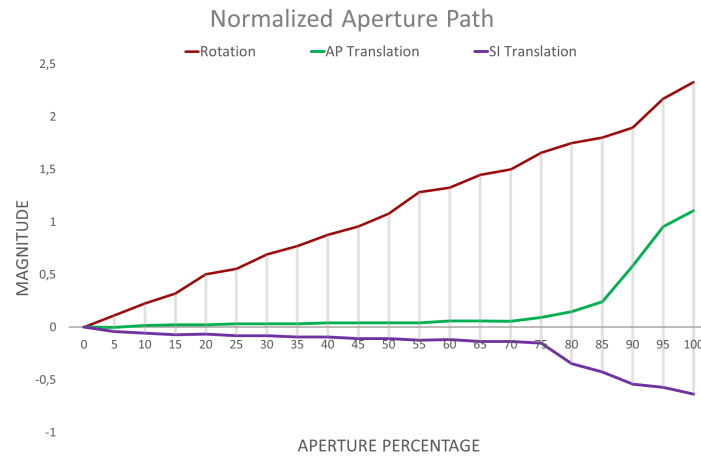
Figure 16: Aperture path graphic showing normalized rotation, antero-posterior and supero-inferior translation movements.

In our proposal, we obtain 3D models of the skull and the jaw for every identification case, either using a 3D scanner or a cone beam computed tomography (CBCT) scan. Then, an expert pinpoints the incisors and the mandibular and cranial condyles in the acquired 3D models during an initialization step. This approach avoids the use of a radiographic analysis for each case tackled. The process of determining this measurement in 3D is thus eased and the resulting error of locating the condylar center is reduced (Fig. 17).

Figure 17: Illustration of the condylar center location and measurement of the mandible radius

The initialization step allows us to: i) position the skull in a standardized anatomic pose (Frankfurt plane) [SM87], ii) automatically obtain the magnitudes describing the mandibular opening movement for a given skull, and iii) directly apply the optimization method proposed in Section 3.3.3. By considering this step, the application of the technique is simplified by not imposing a manual positioning of the skull as in previous automatic and non-automatic SFO methods.
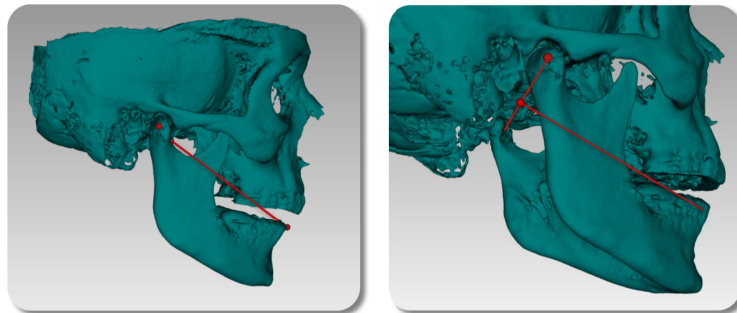
3.3.3 Articulated Genetic Algorithm for Skull-Face Overlay

Our proposal is based on the state-of-the-art genetic SFO method used in [IBC$^+$09, ICDS11, CÁInCA$^+$15, CÁCÁG$^+$17], including the objective function (Eq. I.12). In order to incorporate the mandible articulation by means of the model described in Section 3.3.2.2, the mandible position is also controlled by an optimization parameter. This parameter simulates a different aperture percentage, where 0% corresponds to central occlusion and 100% to the maximum predicted opening in the sagittal plane. This method relocates the mandible landmarks prior to the application of the geometric transformation, adapting their position at every step of the optimization process, until the optimal solution is found. It is worth noticing that the complex SFO problem, which requires the estimation of 12 parameters, is only incremented by a single additional parameter concerning the mandible aperture, achieving a good trade-off between accuracy and optimization complexity.

The extension of the existing real-coded genetic algorithm (RCGA) for automatic SFO [CÁInCA$^+$15] to include the mandible aperture parameter can be achieved in a number of ways. So, in order to look for the best design we have tested three variants, named A1, A2 and A3.

- **A1:** In this variant, the aperture percentage is directly added to the RCGA chromosome and thus undergoes the same evolutionary process as the 12 projection parameters (see Fig. 18). The optimization procedure ensures that the mandible position is estimated together with the skull transformation parameter in each candidate solution during each step of the genetic algorithm. A1 implements the most straightforward approach to extend the RCGA using an additional parameter. One of the advantages of this approach is that the full range of solutions of the extended problem is available. Nevertheless, this approach has an important drawback. The search space is increased and the good performance of the RGCA in optimizing the projection might be hindered by the interaction with the new parameter.
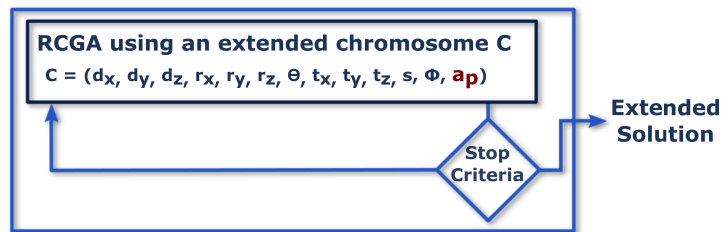
Figure 18: Coding scheme for A1 variant.

- **A2:** A local search to estimate the value of the aperture parameter is performed after the execution of the original RCGA method (see Fig. 19). In this case, the coding scheme and thus the search space size are kept unchanged for the global optimization procedure. Once a final solution is obtained, an exhaustive local search is applied so that all the possible mandible positions predicted by the aperture model are evaluated. It is intended as a refinement stage for the mandible position after the standard automatic SFO procedure, following a problem decomposition approach by means of a genetic local search. A2 is effective solely when the solution provided by the RCGA (i.e. not estimating the aperture) can be turned into a high-quality solution for the extended problem by adjusting only the mandible aperture.

This is likely to happen when the SFO scenario has a small mandible aperture. The advantage of A2 over A1 is that more resources are devoted to optimizing the projection parameters. The inconvenience is that a partial solution is considered at each stage.

Figure 19: Coding scheme for A2 variant.

- **A3:** In this alternative, an exhaustive local search is applied after each generation of the RCGA in order to find the optimal aperture percentage parameter for the best individual found by RCGA, in a similar way to the A2 variant. The genetic evolution is applied to the original 12 unknowns only. Once a generation comes to an end, the opti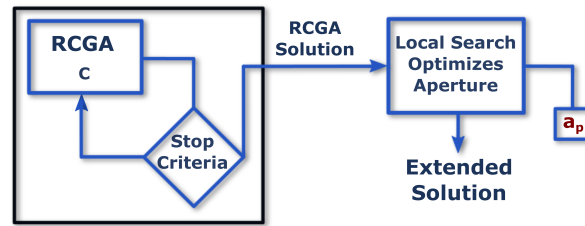mal aperture for the best solution in the population is obtained and the individual is carried out to the next generation modifying its fitness values. A3 is based on a memetic algorithm considering an exhaustive local search applied over a single individual of the population. It thus achieves a compromise between the former two algorithms, by alternating the optimization of the projection and the mandible aperture, which could be more effective than doing it jointly.

Figure 20: Coding scheme for A3 variant.

3.3.4 Ground Truth Data Generation for Skull-Face Overlay

Before presenting and analyzing the results of the experimental study, the procedure to create a ground truth dataset of SFO study cases is described. This dataset will allow us to analyze the performance of our methods, following an objective and reliable procedure.

Previous works in collaboration with forensic experts remark the importance of objective experimental analysis to validate the results and improve the confidence of the CFS technique [IVC$^+$16]. In this proposal we aim to extend the ground truth dataset of [ICCÁ$^+$15] following a similar procedure. To that purpose, we generate a mandible aperture dataset with the collaboration of one subject, which was previously scanned using a CBCT with a neutral facial expression. The subject gave his consent to share his anonymized clinical data for research purposes.

In [ICCÁ$^+$15], 2D photographs for each individual were taken in a pose similar to the mandible pose during the CBCT acquisition, while the authors identified facial pose differences between the two acquisition times as a possible source of error. In this work, this issue is addressed by using simulated AM photos, projected from the actual 3D face model, which ensures a perfect correspondence with that 3D model. Although a single subject was involved, the generation of the data for the experimental study has been quite complex.

The data of a single SFO case consist of a 3D model of the cranium and mandible (both segmented from the CBCT) together with an AM photo showing a specific mandible aperture and overall pose. The photo was simulated by rendering a 3D model of the face with a given mandible aperture in a specific pose (e.g. frontal). To this end, we acquired a series of 3D face models of the subject with different mouth apertures using a hand-held Artec Spider^{TM} 3D scanner. Specifically, the subject was scanned with a 5 *mm* aperture variation (measuring inter-incisor distances). Ten scans were obtained ranging from central occlusion (rest) to the maximum aperture the subject was able to maintain comfortably during the scans (40 *mm*), including a scan of the subject smiling, to let the subject's teeth appear in the 3D model. A gauged resin cast was used to control the aperture distance for each scan (see Fig. 21).

Figure 21: From left to right, the images show the detail of the resin cast used to control fixed aperture distances of 20 *mm*, 40 *mm*, 35 *mm*, and 40 *mm*, respectively.

The process described so far would be enough to create a dataset of SFO cases. However, we still need to compute the (a priori) solution of the cases, i.e., the projection f mapping the skull onto the photo. Note that we simulated the photo by setting the face model in a known pose (i.e., according to a transformation f') but this is not the same f that projects the skull. As the face and skull models are obtained from different sources, they are not aligned. To this end, we performed a manual segmentation of the CBCT to extract an additional face model (which is inherently aligned with the skull) and used it to align the dataset face models. Once the models are aligned, both f and f' are considered identical and provide the optimal solution used as a reference to measure the quality of the results. Finally, an expert marked the homologous cranial and facial 3D landmarks in the different models. Fig. 22 summarizes the data acquisition procedure for validation purposes.
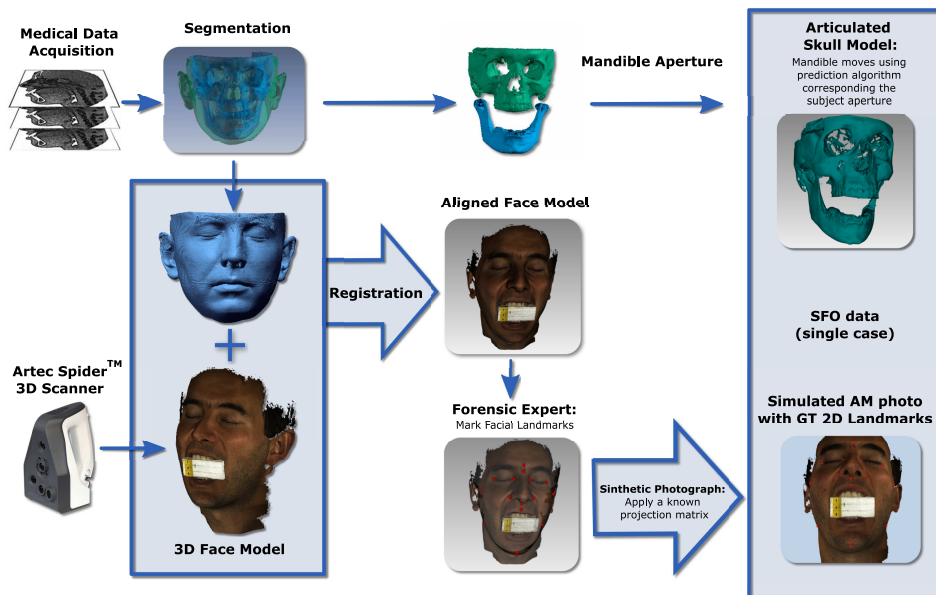

Figure 22: Overview of the Ground Truth data generation procedure

3.3.5 Experiments: Articulated SFO

The objective of this experimental study is two-fold: i) to analyze the performance of the proposed articulated approach in comparison with the state-of-the-art automatic SFO method presented in [CÁInCA⁺15]; and ii) to validate the obtained identification results, which is an essential requirement to verify the reliability and robustness of our new SFO approach.

3.3.5.1 Experimental Setup

For each of the 10 facial scans in the ground truth dataset, three 2D photographs were extracted in different poses (frontal, lateral right, and lateral left) as projections of the 3D model. Thus, the experimental setup consists of 30 SFO instances that provide an objective testbed to analyze and evaluate the performance of our new approach in comparison with the state-of-the-art SFO automatic method, named RCGA [CÁInCA⁺15].

All the experiments have been performed on an Intel Core^TM i7-3537U 2.0 GHz CPU, with 10 GB of RAM, running Windows 10 Professional^TM. For each SFO instance, 30 runs of each algorithm were considered, using a stopping criterion of 600.000 evaluations, and the same parameter configuration that yielded the best results in [CÁInCA⁺15] (see Table I.9). All the proposals that we consider make use of the same fitness function as the RCGA (see Eq. I.12). Regarding the average execution time per run, there was a difference of 5 seconds between the fastest and the slowest algorithm: RCGA (25.05 s) and A1 (30.86 s), respectively.

Parameter	Value	Description
G	600	Generations
P_s	1000	Population size
p_c	0.9	Crossover probability
p_m	0.2	Mutation probability
T_s	2	Tournament size
η	1	SBX Distribution index

Table I.9: RCGA parameter configuration.

3.3.5.2 Analysis of Experimental Results

This section comprises a summarized version of the experimental analysis performed in [BCAV⁺17]. The interested reader is referred to Part II: Section 3 of this dissertation, where the complete study is detailed. Below we analyze the performance of the proposed articulated SFO method in comparison with the current state-of-the-art method, RCGA. Once the skull is overlayed, we first calculate the error (Euclidean distance among landmarks) for all the estimated landmarks and their corresponding ground-truth positions. Then, we compare the mean error of the considered algorithms, which is obtained by averaging each landmark error for every SFO instance.

Table I.10 presents the mean error for each instance, distinguished by case (actual inter-incisor aperture distance) and pose (F= frontal, R = right, and L= left views). Both *Rest* and *Smile* cases correspond to an aperture of approximately 0 mm (closed mouth in central occlusion). The only difference is that, in the latter, the subject is smiling, so his teeth are visible in the 3D model taken as reference.

It is worth noticing that the mean error values obtained are always lower in the frontal instances ($< 1.7\ mm$), while lateral poses tend to increase the resulting error ($> 1.8\ mm$). This behavior is explained by the inherent complexity of the lateral instances, as there are less visible landmarks in the 2D photograph, hindering the optimization process (readers interested in coplanarity problems in image registration are referred to [SCDI09] for a deeper explanation).

Case, Pose	Mean error			
	A1	**A2**	**A3**	**RCGA**
Rest, F	1.54	1.60	**1.52**	1.55
Rest, L	3.23	**3.18**	3.22	3.24
Rest, R	3.20	3.11	3.16	**3.09**
Smile, F	**1.39**	1.40	1.52	1.52
Smile, L	**3.21**	3.28	3.34	3.26
Smile, R	**3.23**	4.33	4.23	4.12
5, F	**1.22**	1.30	1.52	1.30
5, L	3.69	2.89	**2.88**	3.71
5, R	**2.83**	2.93	2.92	2.90
10, F	**1.59**	1.72	2.16	1.86
10, L	**1.88**	2.44	2.32	3.71
10, R	3.48	2.74	**2.51**	2.99
15, F	**1.50**	1.96	2.61	2.54
15, L	**2.44**	3.04	3.14	3.96
15, R	**2.58**	3.11	3.38	4.02
20, F	**1.30**	2.40	2.80	3.27
20, L	**2.91**	3.92	4.02	4.93
20, R	**3.11**	3.65	3.92	4.95
25, F	**1.64**	2.45	3.25	3.47
25, L	**3.74**	4.66	5.01	5.66
25, R	**2.58**	4.79	5.51	5.79
30, F	**1.28**	2.98	3.76	3.84
30, L	**2.51**	4.22	5.69	6.56
30, R	**2.55**	3.83	5.87	6.60
35, F	**1.24**	3.10	3.74	4.04
35, L	**3.71**	4.52	6.23	7.35
35, R	**2.40**	4.66	5.91	7.44
40, F	**1.67**	3.53	4.68	5.01
40, L	**4.57**	6.75	6.90	8.29
40, R	**2.38**	8.14	7.80	9.88

Table I.10:   Mean error for every SFO instance of the ground truth dataset.

Notice also that when the aperture is small (i.e., $0 - 10\ mm$ aperture range), the differences among the results are negligible. This behavior is explained by the fact that the local search variants (A2 and A3) only focus on the aperture percentage parameter. They thus improve the mandible position obtained by the basic algorithm scheme with lower apertures. Nevertheless, they are unable to properly determine the skull location when the aperture is larger. That is the case of all the remaining problem instances, where A1 reduces the resulting error notably (more than $0.45\ mm$). This improvement confirms our hypothesis that A1 would be able to reach more accurate solutions since it explores the whole search space at the expense of increasing the optimization complexity. In view of these results, the inclusion of the aperture percentage parameter within the evolutionary optimization process is strongly recommended.

Fig. 23 provides an overview of the mean error results. Overall, the range of error values for cases having lateral poses is much larger than that of frontal ones. This is also reflected in the corresponding mean values. However, most solutions are actually quite close to the mean, with a few solutions scoring very high errors. In terms of variability, A1 shows the most consistent behavior. In comparison, A2, A3, and RCGA delivered a broader range of results, with A2 being the worst in the group.

In order to complement the analysis of the results, we applied different post-hoc statistical analysis [GH08], which are summarized in Table I.11.   First, we considered Friedman's nonparametric test [Fri40], which aims to test the null hypothesis stating that there is no difference among the resulting mean errors of all the algorithms. The result of applying Friedman's test is $\chi^2_F = 44,52$, and the corresponding p-value is $1,17 \cdot 10^{-09}$. Since the p-value is much lower than the considered level of significance ($\alpha = 0.05$), the test concludes that there are significant differences among the observed results.
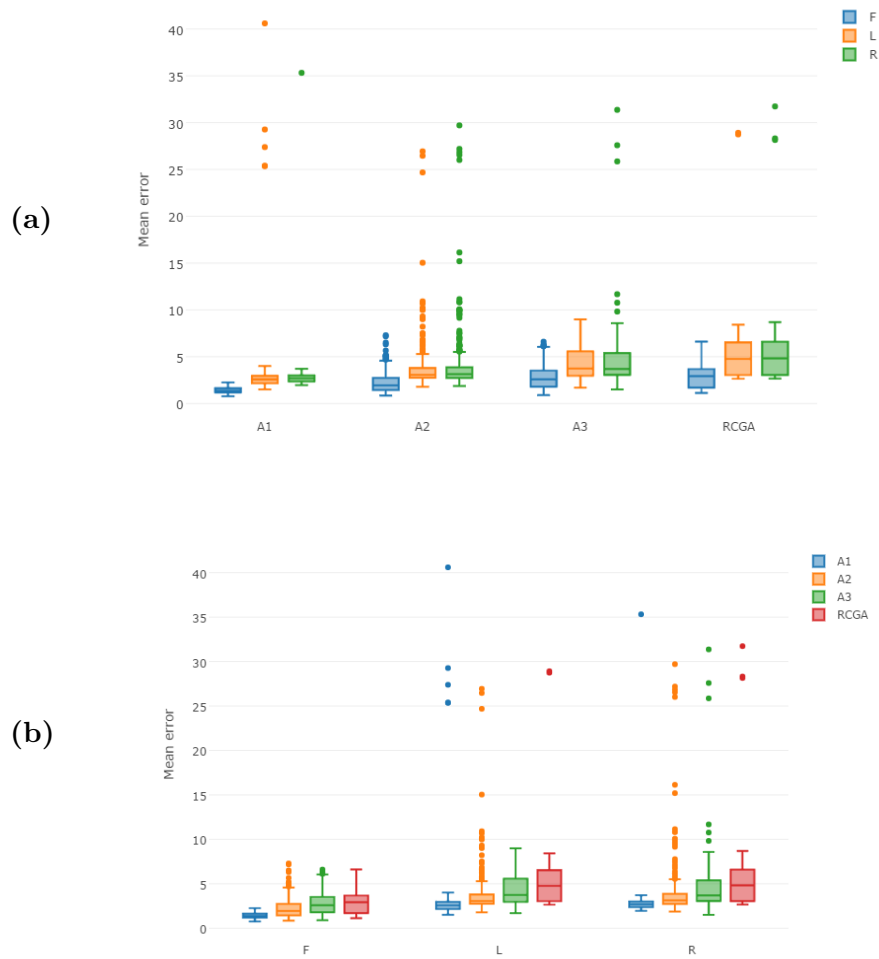
Figure 23: Box-plots corresponding to mean error (in mm) grouped by: (a) algorithm and pose, (b) pose and algorithm.

Algorithm	Mean Error	Ranking	Bonferroni-Dunn p	Holm p
A1	2.49	1.36	-	-
A2	3.42	2.26	0,02	0,02
A3	3.85	2.86	$< 10^{-4}$	$< 10^{-4}$
RCGA	4.36	3.50	$< 10^{-9}$	$< 10^{-9}$

Table I.11: From left to right: Mean error of each optimization algorithm, ranking obtained through Friedman's test, statistical p-values on the results with A1 as control method for Bonferroni-Dunn's test, and for Holm's test.

Next, we considered the Bonferroni-Dunn test [Dun61] to detect significant differences between a control approach and the rest of the algorithms. In this case, the control algorithm is the proposed variant A1 and the results are graphically represented in Fig. 24. A Bonferroni-Dunn's graphic illustrates the significance of the difference among rankings obtained for each algorithm considering the mean error. The horizontal line shows the significance threshold at $\alpha = 0.05$, whose height corresponds to the sum of the ranking of the control method (A1) and the critical difference value computed by the Bonferroni-Dunn method. Thus, algorithms exceeding the threshold line perform significantly worse than the control method for the considered α.

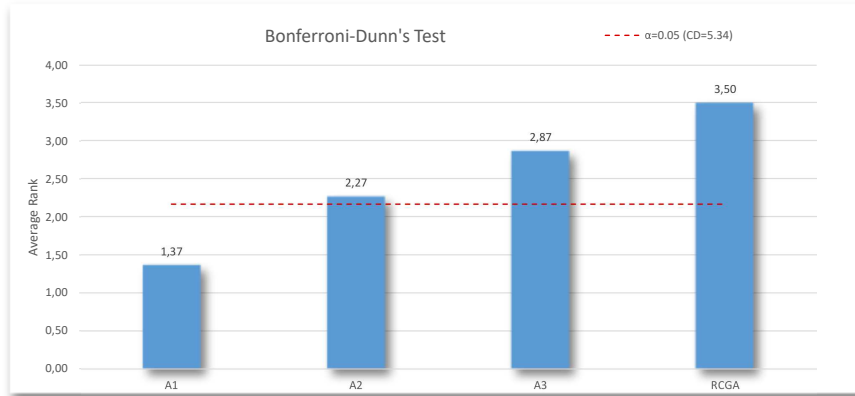Figure 24: Bonferroni-Dunn's graphic corresponding to the results.

The Bonferroni-Dunn's test shows us that the control method A1 is significantly better than every other algorithm at $\alpha = 0.05$ with an ample performance gap. Moreover, we complemented the statistical analysis with another powerful test, such as Holm's, while still comparing A1 with the rest of the methods. The p-value results for Bonferroni-Dunn's and Holm's procedures are included in Table I.11, providing similar results with p-values smaller than $\alpha = 0.05$.

3.3.5.3   Qualitative Analysis of Results

A visual assessment of the quality of the results is depicted in Fig. 25 (including frontal and lateral cases, respectively). Both figures facilitate a comparative visual analysis between the best SFO solution (i.e., the solution achieved by A1 with the lowest mean error in the 30 runs) and the solution obtained by the state-of-the-art RCGA method. Each visual SFO solution consists of the superimposition of the skull and jaw 3D models over the simulated photograph. It also represents the resulting projections of two set of points: cranial landmarks (in green circles) and facial landmarks (in red boxes).

In summary, the visual analysis corroborates the previous numerical analysis, clearly showing the substantial improvement of A1 over RCGA in terms of quality of the results. A1 obtains accurate superimpositions particularly when the mandibular aperture is greater than 10 *mm* (this effect is distinctly visible in the lateral views of Fig. 25). Considering most of the instances, RCGA manages to obtain relatively good mean error values but at the expense of introducing a considerable distortion of the skull (mostly noticeable for bigger apertures, see the last row in Fig. 25).
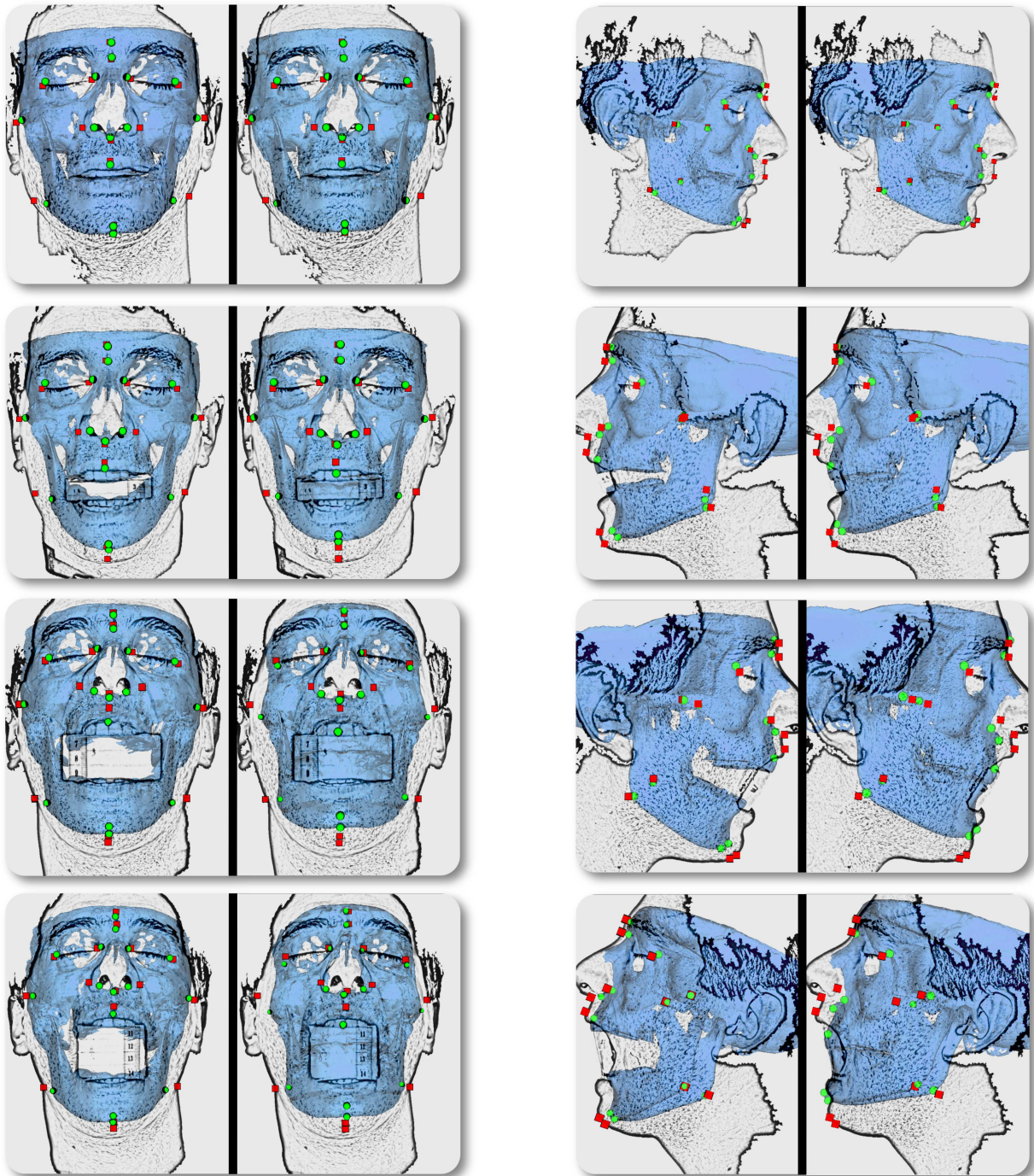
Figure 25: Visual SFO results: A1 (left images) and RCGA (right images) for each of the frontal and lateral instances (Smile, 15, 30, and 40).

4   Summary and Discussion of the Obtained Results

The current section summarizes the main results obtained in this PhD dissertation with a specific subsection for the discussion of each of the main achievements.

4.1   Handling Complexity in Image Registration: Bacterial Foraging Optimization

Regarding the RIR problem, the pipeline proposed to improve range images allowed us to overcome one of the main handicaps of ToF cameras (their low quality) while exploiting their best advantage (high frame rates). The techniques considered for the pipeline not only enhanced substantially the quality of the images but also demonstrated to be computationally efficient for a rapid IP. In addition, the combination of 2D and 3D feature descriptors allowed us to extract meaningful information from both intensity distributions and surface curvature. This approach avoids certain problems present in similar methods when corrupted 3D geometric information is present.

The preprocessing stage facilitated the posterior registration of the ToF range images. This helped the state-of-the-art methods to obtain good, similar results, achieving accurate registrations in most of the problem instances. In addition, several configurations of BFOA variants were able to outperform the results of the state-of-the-art methods, both in terms of accuracy (reaching better minimum results) and robustness (showing a stable averaged performance). Moreover, the two-stage approach proved to be an effective mean to exploit the good results provided by the global optimization algorithms during the prealignment phase.

As for the proposed BFOA variants the conclusions of the designed benchmark were clear, the method proposed by Dasgupta et al. [DDAB09] (A2) highly improved the original BFOA design (A0) [Pas02]. The different components introduced in this work offered alternative means to explore the search space with different outcomes. For example, the variation of the Δ operator showed a negative impact in both minimum and mean results, while elite set and the $BLX - \alpha$ crossover operator improved the accuracy of the variant and the novel elimination-dispersal step improved the exploration capabilities of the algorithm. According to memetic variants, the use of BOBYQA as local search benefited the resulting accuracy, as opposed to those variants integrating XLS, which were more robust alternatives in average. Hence, it is not only a matter of the simple addition of advanced components but a careful design is required.

Concerning the MIR problem, the conditions of the considered scenarios are even more complex than RIR instances and all the compared methods were affected by the different restrictions. Therefore, some algorithms were able to perform well in terms of the objective function, yet failed to provide an acceptable result when the registration was evaluated in terms of the evaluation metric (e.g. PSO-W, Table I.8). The behavior of the proposed BFOA variants was also affected and none of the alternatives was able to outperform the state-of-the-art method in terms of robustness. Nevertheless, two variants obtained the most accurate registration results. As a conclusion derived from the No Free Lunch theorem [WM97], the success of any optimization algorithm depends on the particular context of a problem instance. Even so, BFOA emerged as a competitive SI alternative reaching high quality results in the two challenging problems addressed. In addition, the exhaustive design of alternative components together with the highly modular scheme of the algorithm make BFOA a versatile and adaptable MH suitable to handle complex optimization problems.

4.2   Mandible Articulation in Skull Face Overlay: Genetic Algorithms

The traditional CFS technique requires a forensic expert to position the skull in the same pose as the face in the photograph [Cha93, NFKC91]. In previous computer-based SFO methods, two manual

approaches were followed to approximate the mandible aperture [HIWK15]: i) Before capturing the 3D model, the mandible was normally located relative to the cranium so that the model resembled the facial expression of the photograph under study; and ii) Once the mandible and the cranium were scanned, the 3D models were manually positioned according to the relative aperture in the photograph. Both of these manual approaches lead to important shortcomings and they are even more significant given the multiple comparisons performed in CFS to check the correspondence of one skull to different missing/deceased candidates. Additionally, each individual comparison should involve the analysis of one skull against more than one AM photograph of the same person to significantly increase the reliability and accuracy of the method [ASM94, YIMS95, DWK$^+$15]. Overall, this is a very time-consuming task even using an automatic SFO method.

In our proposal, the expert is only required to pinpoint the incisors, and the mandibular and cranial condyles, a much simpler and faster task to carry out. This task is only performed once, no matter which number of AM photos the skull is compared with. Thereby, our articulated SFO approach avoids entirely the time-consuming and error-prone mandible positioning of the skull, besides adding versatility to the procedure as it can adapt the 3D model pose to different mouth openings in the photographs.

Together with obvious anatomical foundations, numeric, visual, and statistical analysis of the experimental results justified the choice to consider the mandible articulation within the SFO method. In addition, the experiments performed to validate the aperture model showed significant differences between our proposal and the state-of-the-art method. More importantly, our best SFO variant (A1) not only reduced notably the mean error of the original approach but it also obtained good results on the estimation of the actual mandible aperture in most of the instances.

Nevertheless, we identified some limitations in our proposal. We introduced different sources of error that may affect the accuracy of the technique:

- Our main concern is the **suitability of the considered mandible aperture model**. Fig. 26 reflects the limitations of the prediction model used in this work. The figure shows a small antero-posterior displacement between the predicted and the actual position of the mandible. This displacement is increased along with the aperture percentage. Considering small apertures, the prediction error is minimal but it is notably higher when considering bigger apertures. Fortunately, most people do not usually appear in photographs with unnatural expressions (with apertures ranging $> 25\ mm$) in real-world cases. As said, a practical parameterization of the problem was essential as long as the prediction model satisfied a minimum reliability standard. Therefore, the good quality of the results validated the suitability of the articulated SFO approach for a real environment, even considering a simplified aperture model.
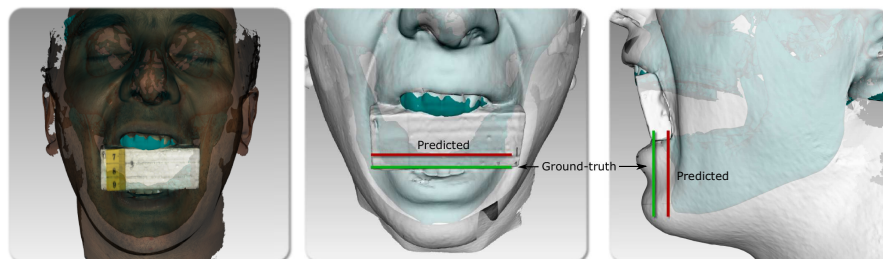

Figure 26: Ground truth position of the aligned face and skull, including the predicted position of the mandible. Red lines indicate the erroneous predicted location of the inferior incisor by the considered model while green lines mark its true location.

- Another source of error is the **manual alignment between 3D scans** to build the ground truth database. In previous works [CÁInCA$^+$15, ICCÁ$^+$15], the 3D models of face and skull were perfectly aligned, as they both were segmented from CBCT scans. In our work, we introduced another source of uncertainty since for each case in the database, the skull model has to be aligned with the facial model. Moreover, each facial model comprises a different amount of soft tissue deformation related to the amount of aperture. We minimized this error by manually registering the CBCT facial model with the undistorted parts for each facial model of the database; thus, it can be considered as imperceptible.

- The SFO approach followed in this work relies on the correspondence between landmarks located in the face and the skull. The ground truth technique introduced here allows us to automatically extract the facial landmarks along with the 2D photographs from the 3D model using a known projection matrix. This technique ensures a perfect correspondence among pairs of facial points in the photograph and the 3D model since the facial pose remains consistent. As the facial models are aligned with the skull model, it is possible to remove potential sources of error regarding landmark placement for the SFO. However, the forensic expert is still required to **pinpoint the landmarks in the 3D models**. The task of manually placing landmarks in 3D models and photographs has already been identified as a significant source of imprecision [CGM$^+$13, CÁIN$^+$14a].

5   Final Conclusions and Future Works

The relevance of registration algorithms for numerous real-world applications is unquestionable, as it is reflected in the extensive research activity of the computer graphics and computer vision communities related to IR. From an optimization viewpoint, the inherent complexity of the IR problems results in a research field in constant evolution, aiming to find reliable and efficient means to provide accurate solutions. Our main motivation is thus to contribute with the identification of different complexity sources in real-world applications and design novel techniques to minimize their impact and overcome the difficulties ahead. Soft computing and AI techniques such as NI or memetic metaheuristics (among other CV techniques) have proven their effectiveness handling ill-conditioned problems in the last few years.

Two main contributions are presented within the framework this PhD dissertation, both adopting recent advances in SC and CV. The first contribution is related to the proposal of a novel and promising SI method, BFOA, to handle two different complex 3D IR problems. As means towards that end, we have conducted a viability study of using video sequences of ToF cameras to tackle 3D reconstruction problems using evolutionary IR techniques. The first element of this study consists of a novel image preprocessing pipeline which provides enhanced range images followed by a subsequent feature extraction procedure which extracts meaningful information combining the intensity distribution (2D) with the geometry (3D) of the scene. The second element refers to the extension of the state of the art in BFOA. Therefore, we have introduced a large number of novel schemes resulting from the combination of different design decisions applied to the original design of the algorithm, obtaining 52 different improved variants. We adapted the different BFOA variants to the 3D reconstruction problem following a pair-wise approach to align adjacent views of the objects. The motivation behind the consideration of this application lies in the complexity derived from a lack of technological means and input data involving noise, deformations, or imprecision.

In addition, we have explored the capabilities of this novel BFOA-based proposal in an alternative application, 3D medical IR. Specifically, we considered a challenging medical problem involving brain MRI scans with different amount of noise and lesions. To overcome the complexity of this problem, a dimensionality reduction has been applied exploiting the curvature of the brain surface and a feature-based approach to perform the registration. Two extensive experimental

studies have been conducted to evaluate the suitability of the proposed BFOA variants in both RIR and MIR applications. From the results obtained in both studies, we remark the high improvement obtained by some of the BFOA variants over the original proposals. Even though the behavior of BFOA does not outperform the current state-of-the-art in some problem instances, the proposed alternative variants and the high modularity of its scheme justify the consideration of BFOA as a valid SI method to handle the complexity of real-world IR problems.

The second contribution is related to the complexity behind a particular 3D/2D IR problem: SFO. Previous works [IBC+09, ICDS11, CÁInCA+15] identified different sources of uncertainty affecting the accuracy of SFO approaches. In this dissertation, we have tackled one such source in order to improve the robustness and reliability of the technique. In particular, we have modeled and integrated the mandible articulation within the SFO optimization algorithm. We have performed a thorough experimental study to analyze the suitability of the proposed articulation model to the SFO technique. Additionally, we have designed a ground-truth database to allow an objective evaluation of the reliability of our proposal. According to the results of this experimentation, the application of a simple mandible aperture model has proven its effectiveness, significantly improving the accuracy and the versatility of the state-of-the-art automatic SFO method. Such an outstanding performance facilitates the use of facial photos where the individual either smiles or opens the mouth partially. Such photos have been usually discarded in real identification scenarios. The availability of new AM pieces of evidence is a crucial consequence of this contribution. Indeed, the use of multiple facial photos of the same individual is essential to increase the reliability of the identification based in CFS [ASM94, YIMS95, DWK+15]

Beyond the achieved improvement concerning the optimization procedure of range and medical IR, several future works can be suggested in an interrelated research line. For example, several NI approaches have been recently proposed exhibiting a strong performance in both search space exploration and exploitation of challenging problems. We believe the study of these novel approaches in different application areas such as IR is beneficial and allows us to explore the extent of these methods. In addition, the last few years have been highly prolific for 3D acquisition technology resulting in a large number of high quality, fast, and portable devices, which also influenced the techniques involved in the reconstruction process. An extensive comparative study of the novel capabilities of these devices is significant for quality demanding applications such as medicine or forensic anthropology.

Regarding the articulated SFO method, the results of the aperture model presented a small error when considering large mandible apertures. Lemoine et al.'s [LXA+07] is an adequate prediction model for the tackled problem since large mandible apertures are not acceptable in real identification cases. We must remark that such a simplified aperture model yielded in huge improvements in both the optimization process and the quality results of the SFO algorithm. Nevertheless, one of our future works aims to analyze the use of more refined and realistic aperture models and study the impact on the reliability of the SFO procedure. In addition, we aim to propose additional innovative designs for the optimization process of the SFO technique.

Summarizing, the results obtained for the three considered applications corroborate the accomplishment of the objectives pursued in this dissertation. The outcomes of the developed research are reflected in three publications in JCR-indexed journals and one conference paper, among other related works. During the development of this PhD thesis, it has been manifested a need of considering a multifaceted approach when tackling interdisciplinary real-world problems. Not only the development of novel mathematical optimization techniques able to handle ill-conditioned scenarios is indispensable, but also the consideration of the whole set of instruments, techniques, and tools provided by diverse research fields.

Part II. Publications

This chapter presents all the scientific papers published. Together they show the work carried out to achieve the stated objectives in this PhD dissertation. Three sections present each of the contribution developed.

1 Quality time-of- flight range imaging for feature-based registration using bacterial foraging

- E. Bermejo, O. Cordón, S. Damas, J. Santamaría. Quality time-of- flight range imaging for feature-based registration using bacterial foraging, Applied Soft Computing, vol. 13, no. 6, pp. 3178–3189, 2013. DOI: 10.1016/j.asoc.2012.08.041.

  - State: Published.
  - Impact Factor (JCR 2013): 2.679.
  - Category: COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE. Order: 20/121. Q1.

Quality time-of-flight range imaging for feature-based registration using bacterial foraging

E. Bermejo[*,a], O. Cordón[a,b,c], S. Damas[b], J. Santamaría[d]

[a]Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain
[b]European Centre for Soft Computing, Mieres, Spain
[c]Centro de Investigación en Tecnologías de La Información y de las Comunicaciones (CITIC-UGR), University of Granada, Spain
[d]Dept. of Computer Science. University of Jaén, Spain

Abstract

Image registration is a widely tackled research topic in the computer vision and the computer graphics fields. This problem aims to find an optimal transformation or correspondence between images acquired under different conditions. Recently, a new 3D image acquisition device based on the time-of-flight technology has appeared which obtains range images from real-time 3D video sequences. In this contribution, we aim to study the feasibility of using this new class of cameras to face the 3D model reconstruction procedure. Our proposal is two-fold. First, we introduce a novel image preprocessing pipeline in order to improve the quality of time-of-flight range images and a subsequent feature extraction method considering both 2D and 3D images. As second major objective, we propose an adaptation of the evolutionary bacterial foraging optimization algorithm, which has recently emerged as a very powerful technique for real parameter optimization and gained a high interest for distributed optimization and control, to tackle the range image registration problem. Finally, we analyse the performance of our proposal against other state-of-the-art evolutionary image registration methods.

Key words: Image registration, Evolutionary computation, 3D modeling, Time of Flight, Bacterial Foraging

1. Introduction

In the last decade, 3D range imaging has been widely used in the community in order to obtain 3D models of real-world objects. In the last few years, range imaging-based technology has opened new lines of research in the computer vision (CV) and the computer graphics (CG) fields [1], mainly focused on the processing of the geometry and the topology of range images, e.g. in noise reduction, analysis, segmentation, classification and 3D model reconstruction. 3D representation of real-world scenarios is essential for a wide variety of applications as robotics, automotive engineering, video surveillance or medical imaging [2–4].

One of the greatest challenges of modern range imaging devices has been the acquisition of quality 3D images in real time. Recently, the time-of-flight (ToF) technology used in some real-time 3D video devices, e.g. the Photonic Mixer Device (PMD) [5], has evolved significantly. Unlike traditional range scanners, the image accuracy of this novel technology is highly dependent on resolution and other ambient and light scene factors. As first approach to face these shortcomings, it is necessary to balance the frame rate in order to reduce the source of error. Nevertheless, more suitable solutions are needed to obtain quality 3D images.

Image registration (IR) [6, 7] is a fundamental task in CV that aims at finding the optimal transformation between two (or more) images. Such transformation estimation is formulated as an optimization problem where the degree of resemblance between images is measured by a similarity metric. The optimization process applied by traditional IR methods, e.g. the Iterative Closest Point algorithm (ICP) [8, 9] is highly influenced by image noise, image

---

discretization, image misalignment, among others. On the other hand, evolutionary computation (EC) [10, 11], in particular evolutionary algorithms (EAs) [4, 12] have demonstrated its ability to overcome some of the shortcomings of traditional IR methods, achieving a robust performance in complex optimization problems.

Recently, natural swarm inspired algorithms like Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO) have found their way into this domain and proved their effectiveness. Following the same trend of swarm-based algorithms, Passino proposed the BFOA in [13]. Application of group foraging strategy of a swarm of E. Coli bacteria in multi-optimal function optimization is the key idea of the new algorithm. Bacteria search for nutrients in a manner to maximize energy obtained per unit time. Individual bacterium also communicates with others by sending signals. A bacterium takes foraging decisions after considering two previous factors. The process, in which a bacterium moves by taking small steps while searching for nutrients, is called chemotaxis and key idea of BFOA is mimicking chemotactic movement of virtual bacteria in the problem search space. It has already drawn the attention of researchers because of its efficiency in solving real-world optimization problems arising in several application domains. Unlike others evolutionary-based approaches, the underlying biology behind the foraging strategy of E. Coli is emulated in an extraordinary manner and used as a simple optimization algorithm. In this work we aim to analyse the performance of a novel design of BFOA facing a challenging real-world problem.

This contribution is structured as follows. Section 2 is aimed to introduce the 3D reconstruction problem using evolutionary computation (EC). Also, this section details the main features of ToF cameras. Next, Section 3 describes our IR proposal based on a specific image processing pipeline using ToF devices and a hybrid EA to tackle the IR problem. Section 4 introduces the experimental study. Finally, Section 5 draws some conclusions and future lines.

2. Background

2.1. Image registration

In the last few years, specialized communities have experienced a growing interest using range scanners for building high-quality 3D models of real-world objects and scenes [3, 14], and avoiding humans to manually produce these models using laborious and error-prone CAD-based approaches [15]. To do so, improved techniques within the 3D model reconstruction pipeline, e.g. IR algorithms, has been contributed to date [2].

There is not a universal design for a hypothetical IR method that could be applicable to all registration tasks, since various considerations on the particular application must be taken into account [6]. However, IR methods usually require the following four components (see Figure 1): two input **Images** named as Scene $I_s = \{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n\}$ and Model $I_m = \{\vec{p}_1', \vec{p}_2', \ldots, \vec{p}_m'\}$, with $\vec{p}_i$ and $\vec{p}_j'$ being image points; a **Registration transformation** f, being a parametric function relating the two images; a **Similarity metric function** F, in order to measure a qualitative value of closeness or degree of fitting between the transformed scene image, noted $f'(I_s)$, and the model image; and an **Optimizer** that looks for the optimal transformation f inside the defined solution search space.

Usually, the iterative closest point (ICP) algorithm [8, 16] is the *de-facto* standard for doing pair-wise IR of range images to build the 3D models in a process called 3D modeling/reconstruction. However, the success of convergence of the algorithm will depend on the initial pose of the images[1].

2.2. Evolutionary image registration

As stated before, the Optimizer component is of great importance in the success of the IR procedure. Since the ICP algorithm was introduced, many contributions have been proposed extending and partially solving its shortcomings [9, 17, 18]. Nevertheless, they still assume it is provided an initial near-optimal alignment of images.

On the other hand, EAs [10, 11] make use of computational models of evolutionary processes as key elements in the design and implementation of computer-based problem solving systems. In the last decade, there is an increasing interest on applying EC principles to complex optimization tasks due to their capability to scape from local optima. Genetic algorithms (GAs) [19, 20] are the more extensively adopted EAs facing optimization problems. These algorithms are based on the mechanism of natural selection, where the search space of the problem is represented as

---

[1]The experiments conducted in [12] demonstrated the poor performance of ICP when facing IR problems considering different degrees of image misalignment.
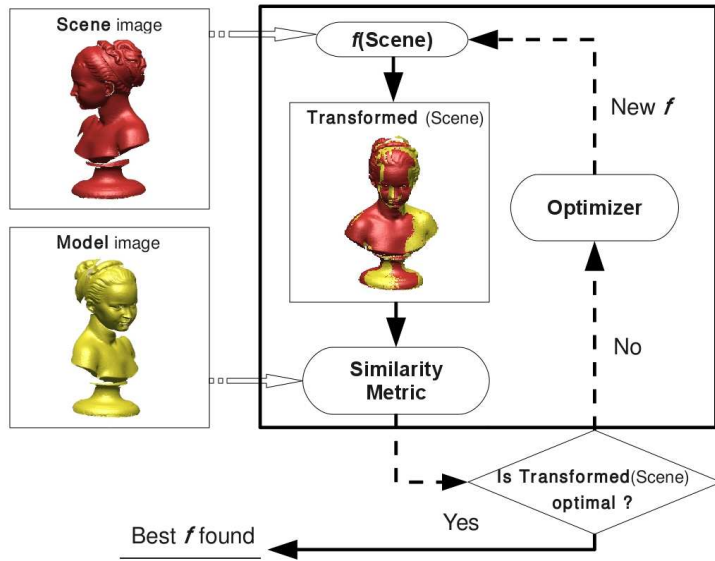
Figure 1: A general framework of the IR optimization process.

a collection of individuals (problem solutions) or chromosomes (conforming a population), each of them simultaneously operating on several points of the search space. At every generation (iteration), some of the candidate solutions are paired and parts of each individual (block of genes) are mixed (crossover operator) to form two new solutions, and additionally, every individual is subject to random changes (mutation operator). The next generation is produced by selecting (selection operator) individuals from the current one on the basis of their fitness or *objective function*, which measures how good is each candidate solution and guides the search space exploration strategy. The objective function is one of the most important components of heuristic methods whose design dramatically affects the performance of the method implemented. Optimization procedures using stochastic schemes as those based on EAs are theoretically and empirically found to provide global near-optimal solutions for complex optimization problems, with several of them from the computer vision and the computer graphics fields [21–25].

The first attempts to face the IR problem using EC can be found in the eighties [26], where a GA [27] was developed for tackling rigid IR of 2D angiography images. Since then, evolutionary IR has become a very active area due to the successful results and several well-known EAs have been considered to tackle the IR optimization process [12].

Unlike ICP-based algorithms, there is no formal proof to asses that EC-based IR methods converge to the global optimum. In order to address this theoretical shortcoming, a two-stage IR approach (see Figure 2) is usually considered in which a first coarser and time consuming stage using EAs, named as pre-alignment, and a refinement step usually[2] applying ICP-based IR algorithms are applied in a serialized fashion. Despite this limitation, evolutionary IR methods are able to achieve accurate results even without requiring any refinement stage as we will see in the experimental section.

2.3. Time of Flight Cameras

ToF cameras are able to capture real-time 3D video. In particular, PMD scanners are based on the time-of-flight principle. Usually, they consist on a CCD or a CMOS sensor and some LEDs or laser diodes that illuminate the scene with a continuous modulated light, which is reflected by the object surface and demodulated in the receiver. The difference between the emitted light and the reflected one generates a phase delay which is used to measure the distances:

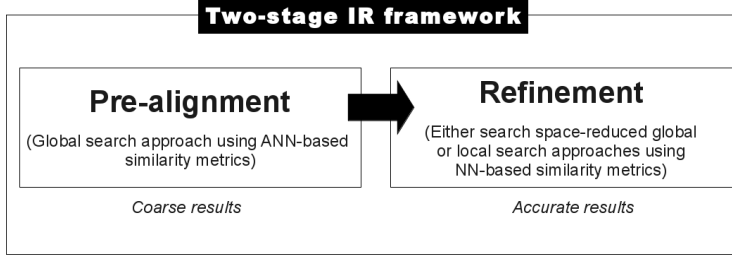[2]Other refinement approaches have been utilized [28].

Figure 2: Pair-wise evolutionary IR approach

$$D_i = \frac{c\varphi_i}{4\pi f_{mod}} \tag{1}$$

where f_{mod} is the modulated signal frequency, c is the light speed, and φ_i is the phase shift of the signal.

Additionally, the phase delay generates an offset from which it is possible to obtain more information about the signal, e.g. intensity or amplitude values. Specifically, the PMD scanner has a 204x204 resolution, 25 fps frame rate, and a CMOS sensor with suppression of background illumination (SBI) features which allows to measure distances between 0.3 and 7 meters. However, one of the main limitations of the real-time 3D video devices is their low resolution [29–32], reaching values between 62x32 and 640x860 pixels. Other difficulties are the appearance of systematic errors as the generation of perfect sinusoidal modulated signal (distance-related error) or by the non-linearity of the electronic components (amplitude-related error). Moreover, non-systematic measurement errors occur when ToF cameras illuminate the whole scene at once. Usually, pixels have a high dynamic range generated by the reflectivity of the object itself, thus leading to measurement errors mostly motivated by cross-reflections or inter-reflections when the signal is reflected on several surfaces (corners, concave objects, etc.).

3. Image Registration Proposal

This section is devoted to introduce both the novel image preprocessing pipeline including feature extraction and the hybrid evolutionary IR algorithm proposed to tackle problem instances considering range images acquired by ToF devices.

3.1. Image Processing Pipeline

As said, it is necessary to reduce the source of noise in order to improve the quality of range images acquired by ToF cameras. Our novel approach uses a three-staged pipeline scheme in which range images obtained from 3D video sequences are enhanced by successively reducing the level of noise. Next, the improved images are used to obtain a small set of robust 3D features (that will guide the IR process) as final step of the procedure.

The first stage faces the noise introduced in the calibration of the emitted signal [33] and consists on the acquisition of an averaged image of the object by means of considering different values of both the integration time and the modulated frequency parameters. Additionally, the frame rate is reduced, which is a quite effective solution as can be shown in Figure 3.

Despite the application of this first noise reduction stage, distance measurements are not accurate enough. Median or amplitude correction filters can be used [34]. Nevertheless, these solutions are prone to provide robust results in presence of remaining noise. Then, the second stage uses an enhancement technique based on the *Shading Constraint* [35] principle. This technique follows the *Shape from Shading* (SfS) [36] approach which is based on the statistical dependence between range (3D) and intensity (2D) information. The code of this stage was provided by the *Action Recognition and Tracking based on Time-of-Flight Sensors* project (ARTTS, www.artts.eu). Figure 4 shows enhanced results using this novel technique when considering the local albedo operation mode [35].
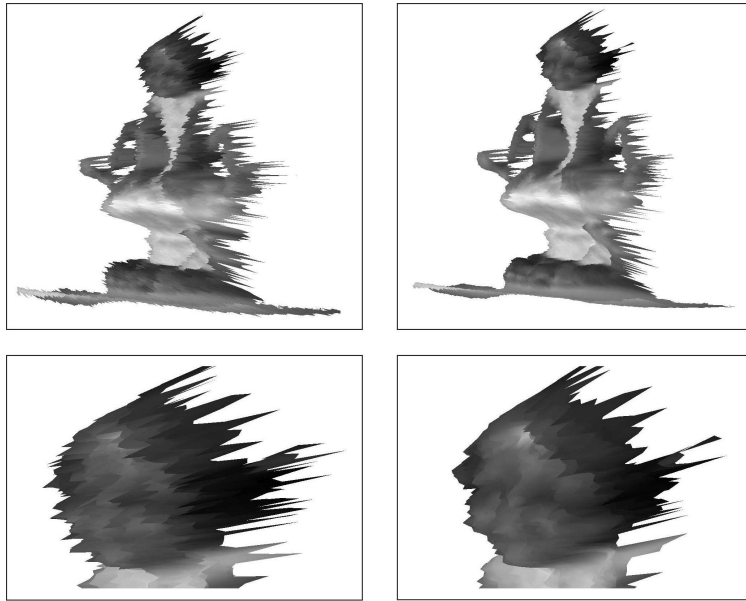
4

Figure 3: Comparison between the original acquired image (left column) and the corresponding image when the first stage is applied (right column).
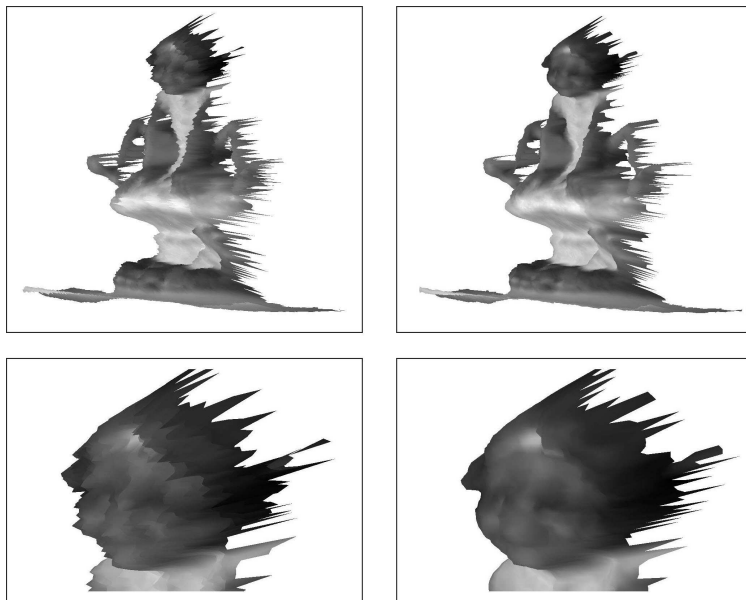

Figure 4: Comparison between the image obtained in the first stage (left column) and the subsequent one provided in the second stage (right column).

Quality improvement of range images achieved following these stages is decisive for the subsequent third stage, i.e. 3D feature extraction. This last step of the proposed pipeline aims to obtain a reduced subset of characteristic range image points in order to both speed up and guide the optimization process of the subsequent IR task. Specifically, it allows to extract two subsets of 3D features as follows:

- First, the method proposed by Castellani et al. [37] is used to extract a small quantity of disperse interest points. This technique builds a geometric descriptor for every salient point which allows to derive information about prominence levels, curvature, and normals. The main advantage of this method is its robustness against noise, holes and occlusions.

- The previous subset of features is increased with additional points taking advantage of the intensity (gray-level) value acquired for each range image point. To do so, we considered the *Speeded Up Robust Features* (SURF) [38] algorithm which is one of the most used invariant feature extraction methods in CV due to its robustness against to noise, errors, and geometric distortions.

Then, our feature extraction proposal exploits the 2D intensity values to extract more relevant information, thus avoiding certain problems shown by similar methods when corrupted 3D geometric information is present. Figure 5 shows a graphical summary of the proposed image preprocessing pipeline.

3.2. Hybrid Bacterial Foraging Algorithm

Recently, Passino has proposed the BFOA search strategy [13]. It is based on the *foraging* approach of the *E. Coli* bacteria, which consists on the following four steps:

Chemotaxis

This process simulates the foraging behaviour of the bacteria according to movement in two different ways. Depending on the food concentration, it swims for a period of time in the same direction or it tumbles to change its current direction. Mathematically, $\theta^i(j, k, l)$ represents the ith bacterium in the jth chemotactic, k^{th} reproductive, and l^{th} elimination and dispersal phases. The tumble movement is modelled as in [13]:

$$\phi(i) = \frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}} \tag{2}$$

where Δ indicates a random vector which direction lies in the range $[-1, 1]$. Thus, the chemotactic step is defined:

$$\theta^i(j + 1, k, l) = \theta^i(j, k, l) + C(i)\phi(i) \tag{3}$$

where $C(i)$ represents the step size taken during the swim step.

Swarming

A social behaviour is simulated in this stage. Those bacteria placed in locations with high amount of nutrients tends to attract other bacteria, while those placed in hazardous zones tend to repel them. This cell to cell attraction-repulsion behaviour is modelled as follows:

$$
\begin{aligned}
J_{cc}(\theta, P(j, k, l)) &= \sum_{i=1}^{S} J_{cc}(\theta, \theta^i(j, k, l)) \\
&= \sum_{i=1}^{S} [-d_{attract} exp(-w_{attract} \Sigma_{m=1}^{p}(\theta_m - \theta_m^i)^2)] \\
&+ \sum_{i=1}^{S} [-h_{repell} exp(-w_{repell} \Sigma_{m=1}^{p}(\theta_m - \theta_m^i)^2)]
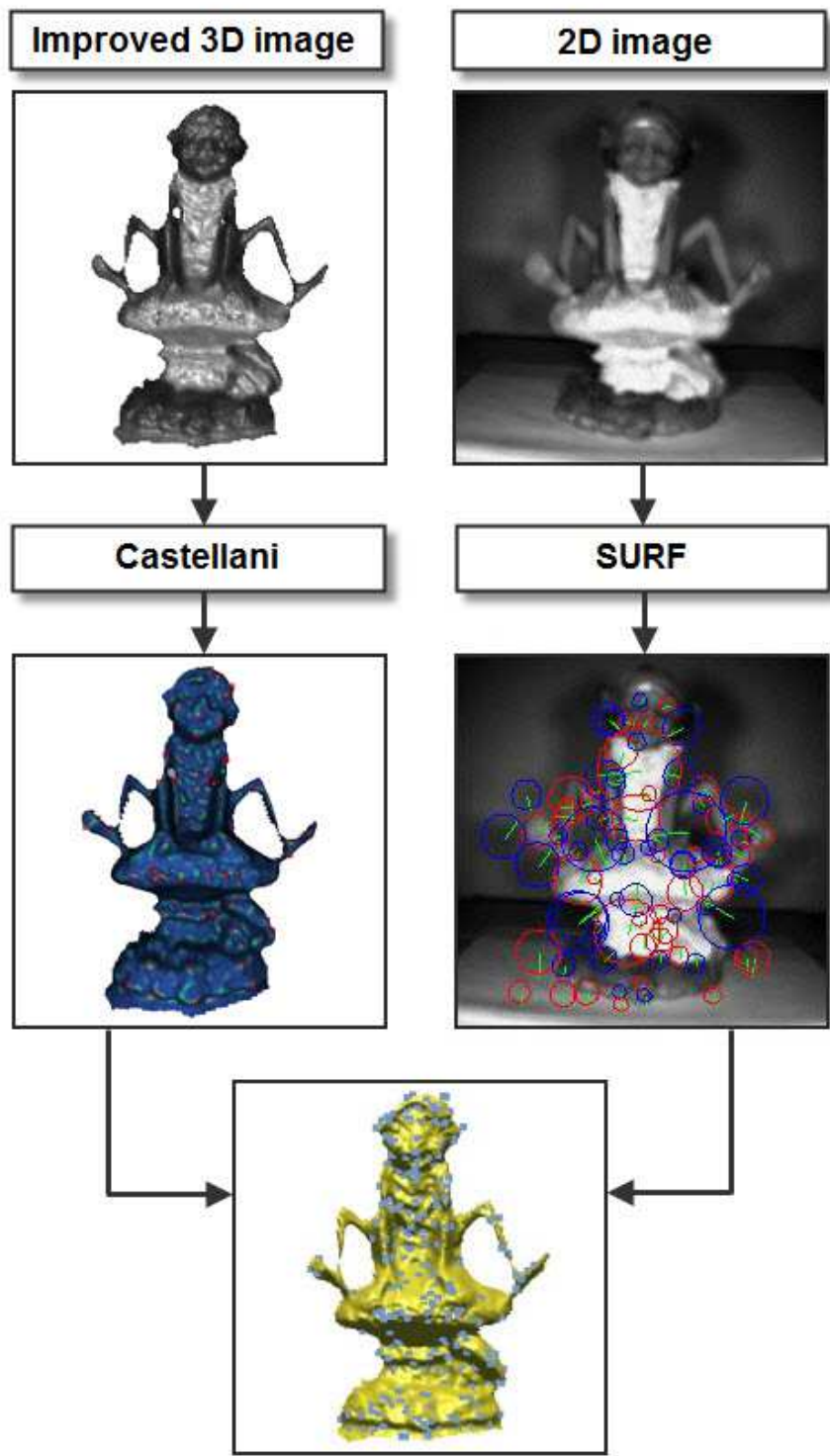\end{aligned}
\tag{4}
$$

6

Figure 5: The proposed image preprocessing pipeline.

where $J_{cc}(\theta, P(j, k, l))$ is the swarming value to be added to the objective function during the chemotactic step j; S is the total number of bacteria; p is the number of variables (or dimensionality of the problem) to be optimized; $d_{attract}, w_{attract}, h_{repell}$, and w_{repell} are different coefficients that represent the attractant and the repellant intensity factors of the bacteria.

Reproduction

The least healthy bacteria, i.e. those who have found less amount of nutrients during the chemotaxis, will die during the reproductive step. On the other hand, the most healthy bacteria will be asexually split in two bacteria. Only the first half of the bacteria will be considered for reproduction by means of replacing the remaining bacteria in order to keep the swarm size constant.

Elimination-Dispersal

The last step consists on the elimination or dispersal of some bacteria based on the simulation of sudden or gradual changes in the local environment, e.g. a rise of the temperature may kill a group of bacteria within a region. This event is simulated by the elimination of some bacteria with a small probability and randomly initializing a new bacteria for replacement.

The pseudo-code of the canonical BFOA [13] is described in the following:

Step 1 Initialize $p, S, N_c, N_s, N_{re}, N_{ed}, P_{ed}, C(i), \theta^i$
   where:
   p: search space dimension
   S: total number of bacterium in the population
   N_c: number of chemotaxis steps
   N_s: maximum number of swims
   N_{re}: number of reproductive steps
   N_{ed}: number of elimination-dispersal steps
   P_{ed}: elimination-dispersal probability
   $C(i)$: chemotaxis step size
   θ^i: bacterium i

Step 2 Elimination-dispersal loop l=l+1:

Step 3 Reproduction loop k=k+1:

Step 4 Chemotaxis loop j=j+1:
   For $i = 1, ..., S$ take the chemotactic step:
   (a) Compute fitness:
      $J(i, j, k, l) = J(i, j, k, l) + J_{cc}(\theta^i(j, k, l), P(j, k, l))$
   (b) Tumble:
      i) Generate a random vector: $\Delta_m(i), m = 1, ..., p$
      ii) Move: $\theta^i(j + 1, k, l) = \theta^i(j, k, l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}}$
      iii) Compute the new fitness value $J(i, j + 1, k, l)$
   (c) Swim:
      Initialize $m = 0$; While $m < N_s$ Do
      - $m = m + 1$
      - If $J(i, j, k, l) < J_{last}$
        update J_{last}
        move: $\theta^i(j + 1, k, l) = \theta^i(j, k, l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}}$
      - Else
        $m = N_s$. This is the end of the while statement and the chemotactic step for bacterium i
   (d) Repeat the process with the next bacterium $i = i + 1$ while $i \neq S$

Step 5 If $j < N_c$ go to Step 4

Step 6 Reproduction:

    (a) For each bacterium, k and l, compute:

$$J^i_{health} = \sum_{j=1}^{N_c+1} J(i, j, k, l) \tag{5}$$

        to represent the total health of the bacterium i along its lifetime. Sort bacteria in ascending order of J_{health}

    (b) The worst S_r bacteria (those with the highest J_{health} values) are replaced with the copies of those bacteria with lower J_{health} values

Step 7 If $k < N_{re}$ go to Step 3

Step 8 Elimination-Dispersal: eliminate the entire pool of bacteria with probability P_{ed}

Step 9 If $l < N_{ed}$ go to Step 2

    Otherwise End

The main shortcoming of the canonical BFOA is the oscillation of the state of the bacteria when they are close to the optimal values, performing a considerable number of unnecessary chemotactic steps. In order to avoid this, Dasgupta et al. [39] proposed a refined variant in which the step size is adapted as follows:

$$C = \frac{|J(\theta)|}{|J(\theta)| + \lambda} = \frac{1}{1 + \lambda/|J(\theta)|} \tag{6}$$

where λ is a positive constant. Thus, when $J(\theta)$ is large the step size is accordingly increased ($C \rightarrow 1$), and it will be decreased once the bacteria approaches the global optima. This adaptive approach outperforms the canonical BFOA as shown in [39]. Our hybrid BFOA (HBFOA) proposal is based on the latter. Additionally, HBFOA incorporates the following novel components to achieve a more suitable performance tackling the IR problem:

- It may occur in the original reproduction step that the bacterium with less accumulated life (J_{health}) is not the best bacterium in the swarm. We propose an alternative based on a probabilistic selection of the best bacteria. Instead of sorting the bacteria in ascending order of their accumulated J_{health} values, a probabilistic filter is applied previous to sorting of the bacteria. In first place, the current life value of each bacteria is normalized considering the life accumulated during the chemotaxis. Then, the best fit bacteria to be reproduced are selected with inverse probability of this normalized value. Finally, the bacteria are sorted and the first half reproduce as in the original method. This strategy increases the importance of those bacteria that obtain best fitness values in the last chemotactic cycles.

- Despite the elimination-dispersal step considers a small probability of application ($1/8$), we tested the performance of the algorithm usually gets down, thus losing the progress achieved in the chemotactic steps. Thus, we have considered the best bacterium found so far to guide the dispersal step in order to minimize the negative impact of killing the best bacterium. Moreover, each transformation parameter has a 30% probability of being similar to the best bacterium.

- Besides, we introduce a set of elite solutions with the M (5) best solutions found. After each chemotactic step, if the algorithm finds a solution, p_{new}, better than the worst elite solution, e_m, a recombination mechanism is applied considering p_{new} and a randomly selected elite solution. Recombination is done by using the $BLX - \alpha$ operator [40]. In case of the new solution, r_{new}, is better than p_{new}, the former will replace both the worst elite solution, e_m, and its parent. Otherwise, p_{new} will replace both e_m and its parent.

- Finally, we have introduced a local search (LS) strategy as a hybridization with the Dasgupta et al.'s BFOA variant. In particular, we considered the crossover-based LS (XLS) [41] method which has obtained promising results in previous works [42]. XLS is applied after the reproduction step.

9

3.3. Coding scheme and objective function

As said, the 3D model reconstruction pipeline involves the application of several pair-wise alignments of two adjacent range images (see Figure 1) in order to obtain the final 3D model of the physical object [3]. Therefore, every pair-wise IR method aims to find the Euclidean motion that brings the *scene* view (I_s) into the best possible alignment with the *model* view (I_m). An Euclidean motion is usually considered based on a 3D rigid transformation (f) determined by six or seven real-coded parameters when using either Euler or axis plus angle representation for rotation, respectively. Specifically, Euler-based rotation matrices suffer from the pitfall of the *gimbal-lock* and specialized literature recommends using the second scheme. Thus, we define the rigid transformation as: a rotation $R = (\theta, Axis_x, Axis_y, Axis_z)$ and a translation $\vec{t} = (t_x, t_y, t_z)$, with θ and $\vec{Axis}$ being the angle and axis of rotation, respectively. Then, the transformed points of the *scene* view are denoted by

$$f(\vec{p_i}) = R(\vec{p_i}) + \vec{t}, \quad i = \{1, \ldots, n\} \tag{7}$$

where n is the number of points of the I_s image. Hence, the pair-wise RIR procedure based on EAs can be formulated as an optimization problem developed to search for the Euclidean transformation[3] f^* achieving the best alignment of both $f(I_s)$ and I_m

$$f^* = \arg\min_{f} \; F(I_s, I_m; f) \;\; s.t. : \; f^*(I_s) \cong I_m \tag{8}$$

according to the Similarity metric, F, being optimized. Among others, the median square error (MedSE) is usually considered as the F function in 3D modeling [43] due to its robustness in presence of outliers (e.g. acquired noisy range images) are present in the RIR process, and it can be formulated as:

$$F(I_s, I_m; f) = MedSE(d_i^2), \quad \forall i = \{1, \ldots, n\} \tag{9}$$

where $MedSE()$ corresponds to the median value of all the squared Euclidean distances, d_i^2, between the transformed scene point, $f(\vec{p_i})$, and its corresponding closest point, $\vec{p'_j}$, in the *model* view I_m, that is:

$$d_i^2 = \|f(\vec{p_i}) - \vec{p'_j}\|^2, \quad j = \{1, \ldots, m\} \tag{10}$$

where m is the number of points of the I_m image. Notice that, it can be said that both the F function and either the fitness or the objective function (see Section 2.2) have the same meaning within the optimization process. In order to speed up the computation of the closest point of every $f(\vec{p_i})$ point, indexing structures as kd-trees [28] or the grid closest point (GCP) transform [44] are often used. Specifically, we have considered the kd-tree scheme together with a random sampling of the cloud points in our later experimental study.

4. Experiments

4.1. Datasets

The range images used in the experiments were obtained from video sequences acquired using a PMD[Vision]® Camcube 2.0 camera considering three different objects. Each scan has been carried out every 45 rotation degrees using a turn table [18]. Figure 6 shows the snapshots of the three considered objects, named: Goblin, Teddy, and Pirate. The left column depicts the obtained range images by using the first two steps of the proposed pipelined image processing scheme (Section 3.1).

Table 1 shows the number of points of both the preprocessed range image (raw data) and the extracted features using the proposed image processing pipeline (Section 3.1). It is remarkable the reduced size of data the IR method should deal with, thereby it will allow to speed up such a procedure.

---

[3]For optimization, any RIR solution is represented as a seven-dimensional real-coded vector $x = \langle \theta, Axis_x, Axis_y, Axis_z, t_x, t_y, t_z \rangle$.
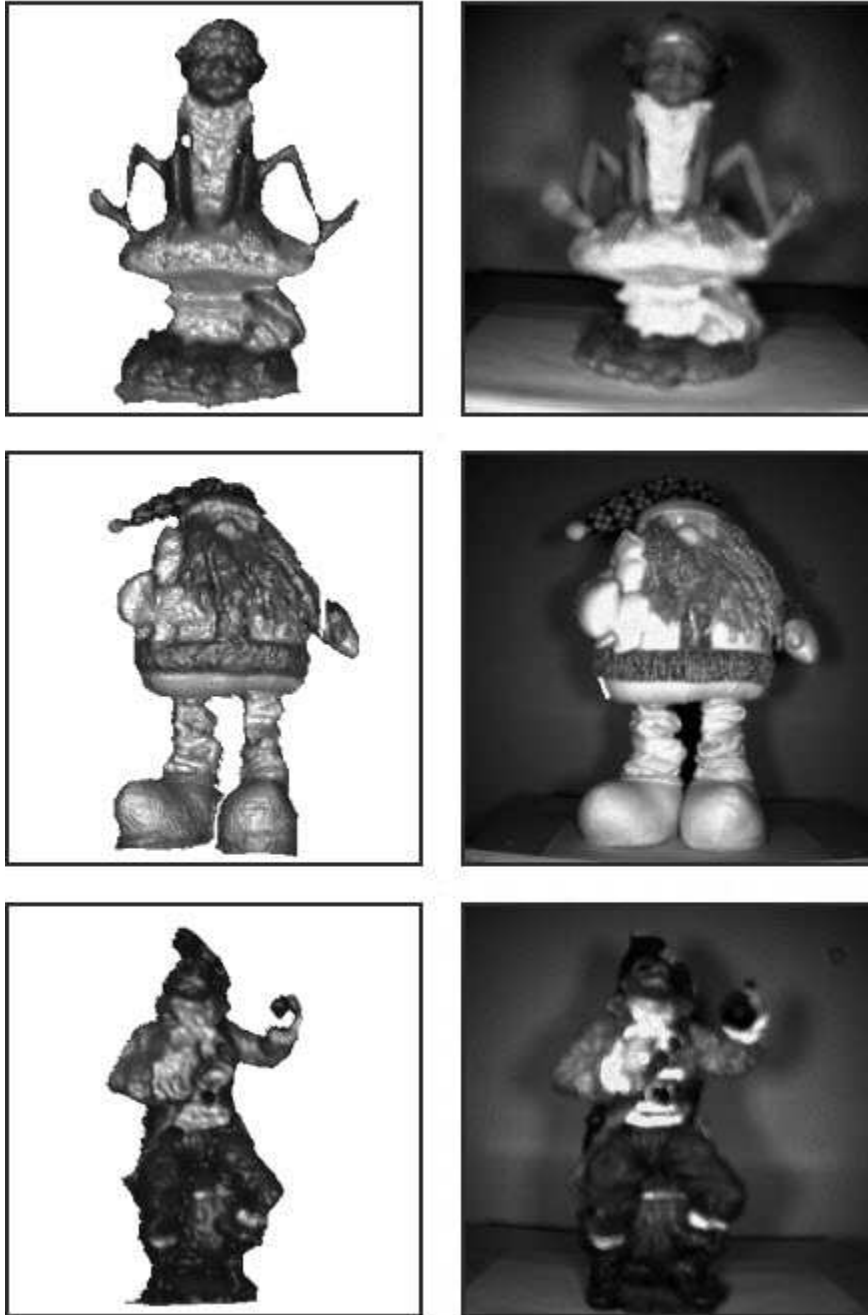
Figure 6: 3D/2D snapshots of the three objects (*Goblin*, *Teddy*, and *Pirate*) acquired using the PMD camera. Textured range images (left) and their corresponding photo (right) are shown.

	Goblin		Teddy		Pirate	
	0°	45°	0°	45°	0°	45°
Raw	11.468	10.983	13.802	13.641	11.340	10.210
Features	224	228	261	259	254	224

Table 1: Number of feature points extracted once the third step of the proposed pipeline is applied.

11

4.2. Experimental Design

Each problem instance to be tackled consists on the registration of two adjacent views (0° and 45°) of every of the three considered objects. We compared the performance of our proposal, HBFOA, against to the Passino's canonical BFOA [13] (Passino) and the Dasgupta et al.'s improved BFOA variant [39] (Dasgupta). Besides, we considered some of the most relevant state-of-the-art IR evolutionary algorithms:

- Luck et al.'s [45] method, which is based on ICP and annealing (ICP-L).

- Chow et al.'s [46] method. It makes use of GAs (GA-C).

- Silva et al.'s [28] method, which makes use of GAs (GA-S).

- Wachowiak et al.'s [47] method. It is based on the particle swarm optimization (PSO) [48] algorithm (PSO-W).

- Santamaría et al.'s [49] method, which is based on the Scatter Search (SS) [50] algorithm (SS-S).

We considered the CPU time as the stop criterion of the prealignment stage (see Section 2.2) in order to perform a fair comparison of all the IR methods. After a preliminary study, we determined that 20 seconds is a suitable value once noticed that every method properly converge to accurate enough solutions. In order to avoid execution dependence, 30 different runs of each IR algorithm have been performed in every problem instance. Additionally, we have set 40 iterations for the refinement step which makes use of the ICP-based algorithm [9].

4.3. Parameter Settings

All the IR methods were run on an Intel® Core™ i7 2.93 GH platform with 4 GB RAM and implemented using C/C++ and the GNU/g++ compiler under the GNU/Linux Ubuntu 10 operating system. For every IR algorithm of the literature, we used the parameter values recommended by the authors in their contributions. Table 2 shows the parameter values considered in Passino, Dasgupta, and HBFOA.

Parameter	Value	Description
n_p	7	Problem dimension.
n_s	16	Swarm size.
N_c	30	Number of chemotactic steps.
N_s	20	Number of swims.
N_{re}	6	Number of reproductive steps.
$Prob_{ed}$	0.125	Elimination-dispersal probability.
N_{ed}	3	Number of elimination-dispersal steps.
C_i	0.1	Step length (fixed only for Passino)
λ	0.0001	Constant value used to calculate the step length (C_i) in Dasgupta and HBFOA algorithms.

Table 2: Parameter settings used in Passino, Dasgupta, and HBFOA algorithms.

4.4. Analysis of Results

Tables 3 and 4 show the results of the prealignment and the refinement stages, respectively. These numerical results refer to the mentioned MedSE metric (Eq. 9) using a Kdtree-based nearest neighbour rule instead of the GCP one considered in the objective function (see Section 3.3).

Regarding the prealignment results, all the BFOA-based IR methods obtain better performance, according to the mean value, than the rest of the methods of the state-of-the-art. Specifically, HBFOA obtains the best mean results in two of the three problem instances (*Goblin* and *Teddy*) and a similar one in the third dataset (*Pirate*) compared with the adaptive BFOA (Dasgupta). According to the minimum values, both HBFOA and Dasgupta IR methods obtain the best results in *Goblin* and *Teddy* instances, respectively. PSO-W achieves the best minimum outcomes in the *Pirate* instance.

	Goblin			Teddy			Pirate		
	m	μ	σ	m	μ	σ	m	μ	σ
Passino	0.00610	0.01211	0.00243	0.01222	0.04671	0.03222	0.00625	0.00853	0.00325
Dasgupta	0.00569	0.01064	0.00383	**0.01182**	0.01594	0.00249	0.00634	**0.00728**	0.00046
HBOFA	**0.00561**	**0.00760**	0.00275	0.01226	**0.01413**	0.00140	0.00647	0.00730	0.00052
ICP-L	0.01167	0.01314	0.00027	0.02788	14.7746	20.8366	0.00746	2.74628	0.50859
GA-C	2.49962	303.279	346.837	11.3260	440.990	498.899	3.29864	412.714	520.820
GA-S	0.00947	0.09527	0.07582	0.01638	0.29109	0.38431	0.00598	0.13806	0.37838
PSO-W	0.00842	0.01208	0.00087	0.01710	0.03689	0.01461	**0.00594**	0.00982	0.00276
SS-S	0.00568	0.00971	0.00340	0.01220	0.01446	0.00161	0.00653	0.00738	0.00039

Table 3: Statistical results obtained in the prealignment stage for each dataset showing minimum (m), mean (μ) and standard deviation (σ) values of MedSE. Results are scaled ($\times 10^3$).

	Goblin			Teddy			Pirate		
	m	μ	σ	m	μ	σ	m	μ	σ
Passino	0.00522	0.00896	0.00151	0.01164	0.02181	0.02115	0.00549	0.00581	0.00012
Dasgupta	0.00519	0.00777	0.00224	**0.01163**	0.01205	0.00014	0.00552	0.00577	0.00011
HBFOA	**0.00518**	**0.00611**	0.00179	0.01169	**0.01198**	0.00015	0.00553	**0.00573**	0.00013
ICP-L	0.00934	0.00978	0.00008	0.01215	0.01284	0.00034	0.00579	0.00589	0.00002
GA-C	0.01296	0.07843	0.02832	0.02851	0.46021	0.23071	0.02499	0.07598	0.03782
GA-S	0.00525	0.06017	0.04673	0.01209	0.09637	0.08075	0.00547	0.03242	0.02784
PSO-W	0.00522	0.00942	0.00112	0.01199	0.01295	0.00144	**0.00546**	0.00586	0.00012
SS-S	**0.00518**	0.00700	0.00180	0.01172	0.01202	0.00015	0.00554	0.00577	0.00011

Table 4: Statistical results obtained in the refinement stage for each dataset showing minimum (m), mean (μ) and standard deviation (σ) values of MedSE. Results are scaled ($\times 10^3$).

Table 4 shows the refinement results. Quite similar outcomes are obtained compared to the prealignment stage, in which HBFOA demonstrates to be the most robust IR method. As shown in [18], all the IR algorithms using the EC paradigm provide accurate enough prealignment solutions that favour the proper convergence of the ICP-based IR method considered for refinement. Specifically, HBFOA obtains the best mean performance in the three problem instances and the best minimum in the *Goblin* dataset. Dasgupta, SS-S, and PSO-W achieve accurate results in the remaining instances.

These results confirm that our proposal, HBFOA, provides a competitive performance compared to the state-of-the-art IR methods. Figure 7 shows the convergence velocity of the evolutionary algorithms in the three problem instances. Figure 8 shows the averaged performance (according to mean value of MedSE) obtained by every algorithm in the three problem instances when applied the prealignment and the refinement stages. This graphic remarks the importance of obtaining an accurate enough prealignment result to promote a suitable convergence of the refinement process. Figure 9 shows the best prealignment and refinement results obtained by HBFOA in every problem instance.

5. Conclusion and Future Works

In this work, we have conducted a viability study of using video sequences of ToF cameras to tackle 3D reconstruction problems using evolutionary IR techniques. Our first contribution consists on a novel image preprocessing pipeline which provides an enhanced range image and a subsequent feature extraction procedure. Specifically, we adopted recent advances in both computer vision and computer graphics fields to develop the latter procedure.

The second major contribution refers to the proposal of an improved variant of the bacterial foraging optimization algorithm for tackling the IR problem. Our hybrid evolutionary IR algorithm has proved to be better than previous bacteria foraging approaches and it also obtains a very competitive performance when compared to the most relevant state-of-the-art IR methods. Thus, we have demonstrated the suitability of using the proposed evolutionary approach to address the 3D reconstruction problem.

Due to the increasing interest of recent market in new ToF devices, we expect the emergence of new cameras offering enhanced features. Specifically, the Microsoft Kinect[TM] is a cutting-edge camera which allows to acquire 3D video images of higher resolution than those obtained by similar scanners such as the PMD. Furthermore, we plan to carry out a comparative study of 3D reconstruction results using different devices based on this recent technology.
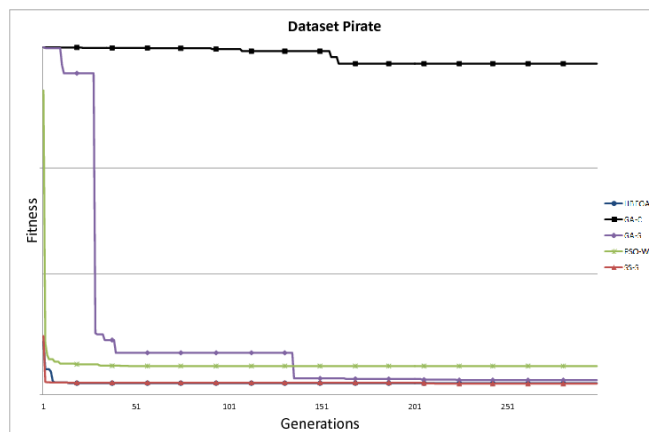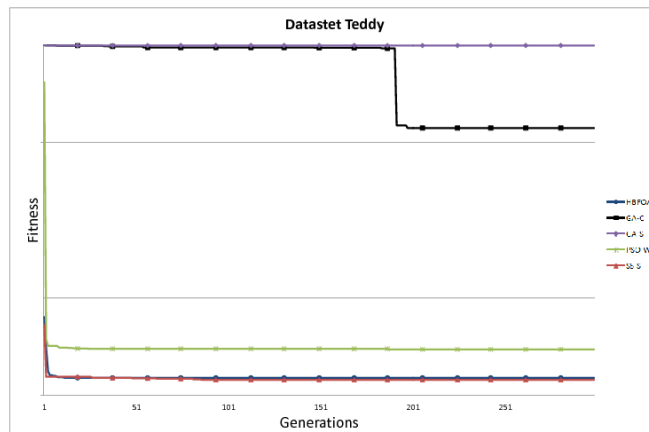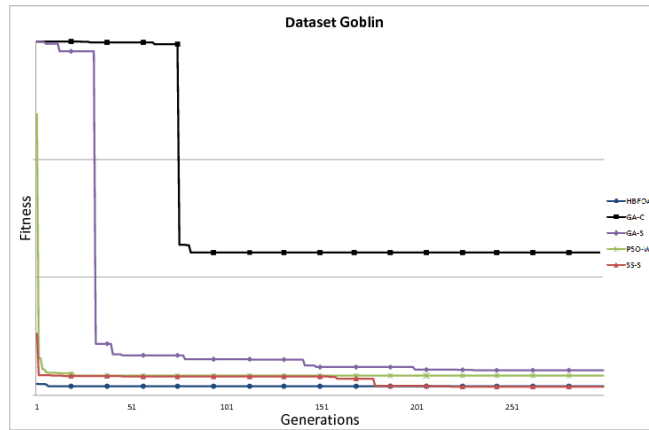
Figure 7: Convergence study of the tested algorithms facing the Goblin, Teddy, and Pirate datasets.
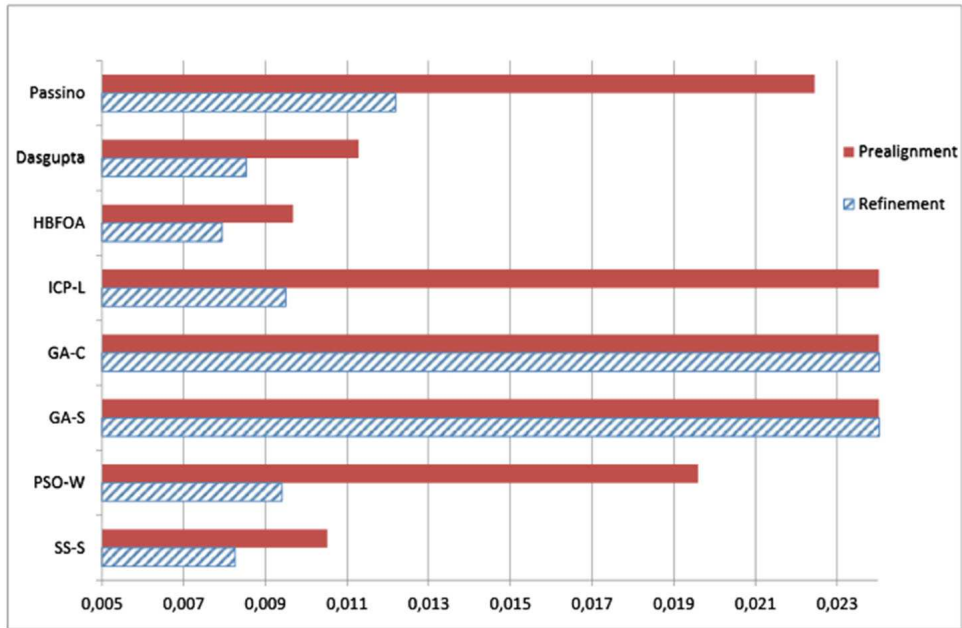
Figure 8: Averaged performance of all the IR methods in the three considered problem instances when applying the prealignment and the refinement stages. Results are scaled ($\times 10^3$).
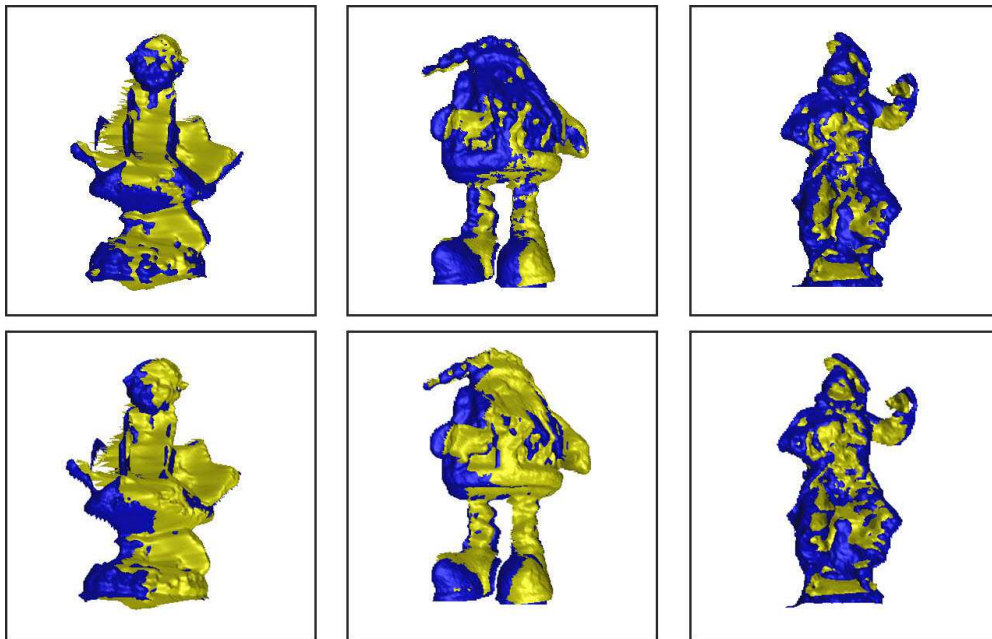
Figure 9: 3D snapshots of the best IR estimations obtained by our HBFOA proposal for the considered datasets *Goblin*, *Teddy*, and *Pirate* after the prealignment (first row) and the refinement (second row) stages.

Acknowledgement

References

[1] R. C. González, R. E. Woods, Digital Image Processing, Prentice-Hall, 2002.

[2] R. J. Campbell, P. J. Flynn, A survey of free-form object representation and recognition techniques, Comput. Vis. Image Underst. 81 (2) (2001) 166–210.

[3] F. Bernardini, H. Rushmeier, The 3D model acquisition pipeline, Computer Graphics Forum 21 (2) (2002) 149–172.

[4] S. Damas, O. Cordón, J. Santamaría, Medical Image Registration Using Evolutionary Computation: A Survey, IEEE Computational Intelligence Magazine 6 (4) (2011) 26–42.

[5] Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, T. Ringbeck, Smart pixel - photonic mixer device (PMD) New system concept of a 3D-imaging camera-on-a-chip, in: 5th Int. Conf. on Mechatronics and Machine Vision in Practice, 259–264, 1998.

[6] B. Zitová, J. Flusser, Image registration methods: a survey, Image Vision Comput. 21 (2003) 977–1000.

[7] J. Salvi, C. Matabosch, D. Fofi, J. Forest, A review of Recent Range Image Registration methods with accuracy evaluation, Image Vision Comput. 25 (5) (2007) 578–596.

[8] P. J. Besl, N. D. McKay, A method for registration of 3D shapes, IEEE T. Pattern Anal. Mach. Intell. 14 (1992) 239–256.

[9] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, Int. J. Comput. Vision 13 (2) (1994) 119–152.

[10] T. Bäck, D. B. Fogel, Z. Michalewicz, Handbook of Evolutionary Computation, IOP Publishing Ltd and Oxford University Press, 1997.

[11] A. Eiben, J. Smith, Introduction to Evolutionary Computation, Springer, Berlin, 2003.

[12] J. Santamaría, O. Cordón, S. Damas, A comparative study of state-of-the-art evolutionary image registration methods for 3D modeling, Comput. Vis. Image Underst. 115 (2011) 1340–1354.

[13] K. Passino, Biomimicry of bacterial foraging for distributed optimization and control, Control Systems Magazine, IEEE 22 (3) (2002) 52–67.

[14] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, D. Fulk, The Digital Michelangelo Project: 3D Scanning of Large Statues, in: ACM SIGGRAPH 2000, 131–144, 2000.

[15] G. Dalley, P. Flynn, Range image registration: A software platform and empirical evaluation, in: Third International Conference on 3-D Digital Imaging and Modeling (3DIM'01), 246–253, 2001.

[16] Y. Chen, G. Medioni, Object modelling by registration of multiple range images, Image Vision Comput. 10 (3) (1992) 145–155.

[17] J. Feldmar, N. Ayache, Rigid, affine and locally affine registration of free-form surfaces, Int. J. Comput. Vision 18 (2) (1996) 99–119.

[18] Y. Liu, Improving ICP with easy implementation for free form surface matching, Pattern Recogn. 37 (2) (2004) 211–226.

[19] D. E. Goldberg, Genetic Algoritms in Search and Optimization, Addison-Wesley, New York, EEUU, 1989.

[20] J. H. Holland, Adaptation in Natural and Artificial Systems, Ann Arbor: The University of Michigan Press, 1975.

[21] Z. H. Hu, Y. S. Ding, W. B. Zhang, Q. Yan, An interactive co-evolutionary CAD system for garment pattern design, Comput.-Aided Des. 40 (12) (2008) 1094–1104.

[22] W. Pang, K. Hui, Interactive evolutionary 3D fractal modeling, Vis. Comput. (2010) 1–17.

[23] X. Qin, Y. Yang, Estimating parameters for procedural texturing by genetic algorithms, Graph. Models 64 (2002) 19–39.

[24] A. Wiens, B. Ross, Gentropy: evolving 2D textures, Comput. Graph. 26 (1) (2002) 75–88.

[25] L. Simon, O. Teboul, P. Koutsourakis, N. Paragios, Random Exploration of the Procedural Space for Single-View 3D Modeling of Buildings, Int. J. Comput. Vision (2010) 1–19.

[26] J. Fitzpatrick, J. Grefenstette, D. Gucht, Image registration by genetic search, in: IEEE Southeast Conference, Louisville, EEUU, 460–464, 1984.

[27] Z. Michalewicz, Genetic algorithms + data structures = evolution programs, Springer-Verlag, 1996.

[28] L. Silva, O. R. P. Bellon, K. L. Boyer, Precision range image registration using a robust surface interpenetration measure and enhanced genetic algorithms, IEEE T. Pattern Anal. Mach. Intell. 27 (5) (2005) 762–776.

[29] R. Lange, 3D Time-of-flight distance measurement with with custom custom solid-state solid-state image sensors in CMOS / CCD-technology, 2000.

[30] B. Büttgen, T. Oggier, M. Lehmann, R. Kaufmann, F. Lustenberger, CCD/CMOS lock-in pixel for range imaging: Challenges, limitations and state-of-the-art, 1st range imaging research day (2005) 21–32.

[31] S. Fuchs, S. May, Calibration and registration for precise surface reconstruction with ToF cameras, in: Proceedings of the Dynamic 3D Imaging Workshop in Conjunction with DAGM (Dyn3D, vol. 1, Citeseer, 2007.

[32] S. May, 3D Time-of-Flight Ranging for Robotic Perception in Dynamic Environments, Ph.D. thesis, Universität Osnabrück, 2009.

[33] S. Fuchs, G. Hirzinger, Extrinsic and depth calibration of ToF-cameras, 2008 IEEE Conference on Computer Vision and Pattern Recognition (2008) 1–6.

[34] S. Oprisescu, D. Falie, M. Ciuc, V. Buzuloiu, Measurements with ToE Cameras and Their Necessary Corrections, 2007 International Symposium on Signals, Circuits and Systems (July) (2007) 1–4.

[35] M. Böhme, M. Haker, T. Martinetz, E. Barth, Shading constraint improves accuracy of time-of-flight measurements, Computer Vision and Image Understanding 114 (12) (2010) 1329–1335.

[36] J. E. Cryer, M. Shah, Shape-from-shading: a survey, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (8) (1999) 690–706.

[37] U. Castellani, M. Cristani, S. Fantoni, V. Murino, Sparse points matching by combining 3D mesh saliency with statistical descriptors, Computer Graphics Forum 27 (2) (2008) 643–652.

[38] H. Bay, a. Ess, T. Tuytelaars, L. Vangool, Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding 110 (3) (2008) 346–359.

[39] S. Dasgupta, S. Das, A. Abraham, A. Biswas, Adaptive Computational Chemotaxis in Bacterial Foraging Optimization: An Analysis, IEEE Transactions on Evolutionary Computation 13 (4) (2009) 919–941.

[40] L. J. Eshelman, Real-coded genetic algorithms and interval schemata, in: L. D. Whitley (Ed.), Foundations of Genetic Algorithms 2, Morgan Kaufmann, San Mateo, EEUU, 187–202, 1993.

[41] H. G. Beyer, K. Deb, On self-adaptive features in real-parameter evolutionary algorithms, IEEE T. Evolut. Comput. 5 (3) (2001) 250–270.

[42] J. Santamaría, O. Cordón, S. Damas, J. García-Torres, A. Quirin, Performance evaluation of memetic approaches in 3D reconstruction of forensic objects, Soft Comput. 13 (8-9) (2009) 883–904.

[43] M. Rodrigues, R. Fisher, Y. Liu, Special issue on registration and fusion of range images, Comput. Vis. Image Underst. 87 (1-3) (2002) 1–7.

[44] S. M. Yamany, M. N. Ahmed, A. A. Farag, A new genetic-based technique for matching 3D curves and surfaces, Pattern Recogn. 32 (1999) 1817–1820.

[45] J. P. Luck, C. Q. Little, W. Hoff, Registration of Range Data Using a Hybrid Simulated Annealing and Iterative Closest Point Algorithm., in: IEEE International Conference on Robotics and Automation (ICRA'00), 3739–3744, 2000.

[46] C. K. Chow, H. T. Tsui, T. Lee, Surface registration using a dynamic genetic algorithm, Pattern Recogn. 37 (2004) 105–117.

[47] M. P. Wachowiak, R. Smolikova, Y. Zheng, J. M. Zurada, A. S. El-Maghraby, An approach to multimodal biomedical image registration utilizing particle swarm optimization, IEEE T. Evolut. Comput. 8 (3) (2004) 289–301.

[48] J. Kennedy, R. Eberhart, Particle swarm optimization, in: IEEE International Conference on Neural Networks, vol. 4, 1942–1948, 1995.

[49] J. Santamaría, O. Cordón, S. Damas, I. Alemán, M. Botella, A Scatter Search-based technique for pair-wise 3D range image registration in forensic anthropology, Soft Comput. 11 (2007) 819–828.

[50] M. Laguna, R. Martí, Scatter search: methodology and implementations in C, Kluwer Academic Publishers, Boston, 2003.

2   A comparative study on the application of advanced bacterial foraging models to image registration

- E. Bermejo, O. Cordón, S. Damas, J. Santamaría. A comparative study on the application of advanced bacterial foraging models to image registration, Information Sciences, vol. 295, pp. 160-181, 2015. DOI: 10.1016/j.ins.2014.10.018.

  - State: Published.
  - Impact Factor (JCR 2015): 3.364.
  - Category: COMPUTER SCIENCE, INFORMATION SYSTEMS. Order: 8/144. Q1.

A comparative study on the application of advanced bacterial foraging models to image registration

E. Bermejo[*,a], O. Cordón[a,b], S. Damas[b], J. Santamaría[c]

[a]*Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain*
[b]*European Centre for Soft Computing, Mieres, Spain*
[c]*Dept. of Computer Science. University of Jaén, Spain*

Abstract

New swarm intelligence approaches as the bacterial foraging optimization algorithm (BFOA) have recently awakened a growing interest in the evolutionary computation community. This fact is due to the promising results obtained by different variants of the latter optimization method in many real-world applications. In this work we aim to take a step ahead in the development of the discipline by introducing a large amount of BFOA variants resulting from the combination of some advanced design decisions applied on Dasgupta et al.'s self-adaptive version. Our goal is thus to achieve an improved understanding of the good properties that the BFOA approach has shown in previous contributions. We will perform an extensive experimental study considering a plethora of algorithmic variants to solve a real-world problem, image registration, a well-known and complex task in computer vision. In particular, more than fifty variants are proposed and tested tackling pair-wise image registration problem instances from two different domains, namely range image registration for 3D model reconstruction and 3D medical image registration. The reported results reveal that BFOA is a versatile approach able to provide very competitive outcomes to face challenging real-world image registration problems when compared to the state-of-the-art evolutionary approaches in the field.

Key words: Swarm intelligence, Bacterial foraging, Image registration, 3D model reconstruction, Time of flight, Medical Imaging

1. Introduction

During the last two decades, natural swarm inspired algorithms like Particle Swarm Optimization (PSO) or Ant Colony Optimization (ACO) have proven their effectiveness in the fields of Computational Intelligence and Meta-heuristics [27]. Following the same trend of swarm-based algorithms, Passino proposed the Bacterial Foraging Optimization Algorithm (BFOA) in [50]. The key idea of this novel algorithm is the application of the foraging strategy of the *Escherichia Coli* bacteria in multi-optimal function optimization. The main objective of this bacteria foraging strategy is to maximize the energy obtained per unit time while searching for nutrients. Individual bacteria also communicate with others by sending signals. A bacterium takes foraging decisions after considering the two latter factors. The process, in which a bacterium moves by taking small steps while searching for nutrients, is called chemotaxis and the key idea of BFOA is mimicking the chemotactic movement of virtual bacteria in the problem search space. This Swarm Intelligence method has already drawn the attention of researchers because of its efficiency when solving real-world optimization problems arising in several application domains [14, 37]. Unlike the inspiration source followed by other evolutionary-based approaches, the underlying biology behind the foraging strategy of E. Coli is emulated in an extraordinary manner and used as a simple optimization algorithm.

BFOA methods have been applied to many real world problems and proved their effectiveness over many variants of Genetic Algorithms (GA) and PSO [44]. For example, Dasgupta et al. proposed a pattern recognition application for the automatic detection of circular shapes based on BFOA [24]. Acharya et al. proposed a BFOA-based Independent Component Analysis [1] to find a linear representation of non-gaussian data where the components are statistically independent. Wan et al. presented a clustering algorithm based on the mechanism analysis of BFOA [67], which outperforms similar algorithms based on ACO and PSO. Besides, El-Abd performed a comparison between some foraging algorithms, including BFOA, and different evolutionary algorithms using a set of benchmarking functions [26].

Meanwhile, image registration (IR) [56, 72] is a fundamental task in CV that aims at finding the optimal transformation between two (or more) images acquired under different conditions: using different sensors, at different time instances, from different viewpoints, or a combination of the latter situations. Such transformation estimation is formulated as an optimization problem where the degree of resemblance between images is measured by a similarity metric. The local optimization process applied by traditional IR methods, e.g. the Iterative Closest Point (ICP) algorithm [7, 71], is highly influenced by image noise, discretization, and misalignment, among others. Nevertheless, evolutionary algorithms (EAs) [2, 25] have demonstrated their ability to overcome some of the shortcomings of traditional IR methods, achieving a robust performance in some complex IR domains [21, 57].

In order to establish a better comprehension of the BFOA optimization properties, we propose a large amount of advanced variants and study their performance to deal with IR, which is suitable for applying EAs and other real-coded metaheuristics. Thereby, we introduce and combine a set of advanced design decisions with the purpose of improving the performance of the standard BFOA proposals in the literature, specifically the self-adaptive version proposed by Dasgupta et. al. [23].

Our aim is to test the different BFOA configurations performance in two different application domains, range image registration (RIR) for 3D model reconstruction, and 3D medical image registration (MIR), thus providing a general view of its behaviour in these fields. To do so, we provide a broad experimentation comparing the proposed BFOA variants performance to the classical IR method, ICP, and other evolutionary and metaheuristic IR methods of the state-of-the-art.

This contribution is structured as follows. Section 2 is aimed to introduce the basis of BFOA and the two IR problem domains considered, 3D model reconstruction from range images and 3D MIR. Next, Section 3 describes the application of the different BFOA variants to tackle the latter two IR problems. Section 4 introduces the developed experimental study for RIR applications, while Section 5 presents the experimental study for MIR applications. Section 6 provides a combined analysis of the behaviour of BFOA variants for both applications. Finally, Section 7 draws some conclusions and future lines.

2. Background

2.1. The bacterial foraging algorithm

The pseudocode of BFOA is described in Algorithm 1, as proposed in [50]. The method consists of the following four steps:

Chemotaxis

This process simulates the foraging behaviour of the bacteria according to movements in two different ways. Depending on the food concentration, the bacterium will swim in the same direction or it will tumble to change its current direction for a period of time. Mathematically, $\theta^i(j, k, l)$ represents the ith bacterium in the jth chemotactic, k^{th} reproductive, and l^{th} elimination and dispersal phases. The tumble movement is modelled as in [50]:

$$\phi(i) = \frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}} \tag{1}$$

where Δ indicates a random vector whose direction lies in the range $[-1, 1]$. Thus, the chemotactic step is defined as:

$$\theta^i(j + 1, k, l) = \theta^i(j, k, l) + C(i)\phi(i) \tag{2}$$

where $C(i)$ represents the step size taken during the swim step.

2

Begin BFOA
1   Initialize parameters
2   **For** l=1 **to** N_{ed} **Do** /*Elimination-dispersal loop*/
3   **For** k=1 **to** N_{re} **Do** /*Reproduction loop*/
4   **For** j=1 **to** N_c **Do** /*Chemotaxis loop*/
5   **For** i=1 **to** S **Do** /*Chemotactic step*/
6   /*Tumble*/
7                   $J(i, j, k, l) \leftarrow J(i, j, k, l) + J_{cc}(\theta^i(j, k, l), P(j, k, l))$
                    /*Compute fitness*/
8                   $\Delta_m(i), m = 1, ..., p$ /*Generate a random direction*/
9                   $\theta^i(j + 1, k, l) \leftarrow \theta^i(j, k, l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}}$
                    /*move bacterium*/
10   evaluate $J(i, j + 1, k, l)$
11   /*Swim*/
12                  $m \leftarrow 0$
13   **While** $m < N_s$ **Do**
14                      $m \leftarrow m + 1$
15   **If** $J(i, j, k, l) < J_{last}$
16   update J_{last}
17                          $\theta^i(j + 1, k, l) \leftarrow \theta^i(j, k, l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta(i)^T \Delta(i)}}$
                            /*move bacterium*/
18   **Else**
19                          $m \leftarrow N_s$
20   **End-If**
21   **End-While**
22   **End-i-For**
23   **End-j-For** /*End Chemotaxis*/
24   **For** i=1 **to** S **Do** /*Reproduction step*/
25              $J_{health}^i \leftarrow \sum_{j=1}^{N_c+1} J(i, j, k, l)$
26   **End-i-For**
27   Sort bacteria in ascending order of J_{health}
28   Duplicate the best S_r bacteria (lower J_{health})
            replacing the worst ones
29   **End-k-For** /*End Reproduction*/
30   **For** i=1 **to** S **Do** /*Elimination-Dispersal step*/
31   **If** $rand() < P_{ed}$
32   Eliminate bacterium and initialize randomly its replacement
33   **End-If**
34   **End-i-For**
35 **End-l-For** /*End Elimination-Dispersal*/
End BFOA

Algorithm 1: Pseudo-code of BFOA.

Swarming

A social behaviour is simulated in this stage. Those bacteria placed in locations with high amount of nutrients tend to attract other bacteria, while those placed in hazardous zones tend to repel them. This cell to cell attraction-repulsion behaviour is modelled as follows:

$$J_{cc}(\theta, P(j, k, l)) = \sum_{i=1}^{S} J_{cc}(\theta, \theta^i(j, k, l)) = \tag{3}$$

$$= \sum_{i=1}^{S}[-d_{attract} exp(-w_{attract}\Sigma_{m=1}^{p}(\theta_m - \theta_m^i)^2)]$$

$$+ \sum_{i=1}^{S}[-h_{repell} exp(-w_{repell}\Sigma_{m=1}^{p}(\theta_m - \theta_m^i)^2)]$$

where $J_{cc}(\theta, P(j, k, l))$ is the swarming value to be added to the objective function during the chemotactic step j; S is the total number of bacteria; p is the number of variables (or problem dimension) to be optimized; $d_{attract}$, $w_{attract}$, h_{repell}, and w_{repell} are different coefficients that represent the attractant and the repellant intensity factors of the bacteria.

Reproduction

The least healthy bacteria, i.e. those who have found less amount of nutrients during the chemotaxis, will die during the reproductive step. On the opposite, the most healthy bacteria will be asexually split into two bacteria. Only the first half of the bacteria will be considered for reproduction by replacing the remaining bacteria in order to keep the swarm size constant.

Elimination-Dispersal

The last step consists of the elimination or dispersal of some bacteria based on the simulation of sudden or gradual changes in the local environment, e.g. a rise of the temperature may kill a group of bacteria within a region. This event is simulated by the elimination of some bacteria with a small probability and randomly initializing the new bacteria for replacement.

The main shortcoming of this canonical scheme of BFOA is the oscillation of the bacteria's state when they are close to the optimal values. A considerable number of unnecessary chemotactic steps are usually performed. In order to avoid this, Dasgupta et al. [23] proposed a refined variant in which the step size is adapted as follows:

$$C = \frac{|J(\theta)|}{|J(\theta)| + \lambda} = \frac{1}{1 + \lambda/|J(\theta)|} \tag{4}$$

where λ is a positive constant. Thus, when $J(\theta)$ is large, the step size is accordingly increased ($C \to 1$), and it will be decreased once the bacteria approaches the global optima. This adaptive approach outperforms the canonical BFOA as shown in [22, 23].

2.2. Image registration

Image registration is an important research field in digital image processing [32]. IR applications cover a broad range of real-world problems, such as remote sensing, artificial vision, computer-aided design (CAD), or medical imaging. The key idea of the IR process is focused on determining the unknown parametric transformation that relates two images by placing them in a common coordinate system, bringing corresponding points as close as possible.

There is not a universal design for a hypothetical IR method that could be applicable to all registration tasks, since various considerations on the particular application must be taken into account [72]. However, IR methods usually require the following four components (see Figure 1): two input **Images** named as Scene $I_s = \{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n\}$ and Model $I_m = \{\vec{p}_1', \vec{p}_2', \ldots, \vec{p}_m'\}$, with $\vec{p}_i$ and $\vec{p}_j'$ being image points; a **Registration transformation** f, being a parametric function relating the two images; a **Similarity metric function** F, in order to measure a qualitative value of closeness or degree of fitting between the transformed scene image, noted $f'(I_s)$, and the model image; and an **Optimizer** that looks for the optimal transformation f inside the defined solution search space.

Usually, the ICP algorithm [7, 15] is the *de-facto* standard for doing a pair-wise registration of two images. It is a well known tool in CAD and medical imaging applications. ICP is an iterative gradient-based method based on the least squares estimation of the IR transformation from the computed matching between scene and model points considering the closest assignment rule. Thus, it is not guided by the similarity metric, but by the computed matching. The original ICP proposal has three main drawbacks [7]: i) the success of convergence of the algorithm will depend on the initial image pose and it likely gets trapped in local optima solutions[1], which forces the user to manually assist the procedure to overcome these situations; ii) one of the two images (usually the scene view) should be contained in the other, e.g. in feature-based IR problems, the geometric primitives of one image should be a subset of those in its counterpart image; iii) it can only handle normally distributed observations, as it is susceptible to gross statistical outliers. For those reasons, traditional IR methods are usually prone to be trapped in local minima [41, 43, 55, 60], providing incorrect registration transformation estimations.

---

[1]The experiments conducted in [21] and [57] demonstrated the poor performance of ICP when facing IR problems considering different degrees of image misalignment.
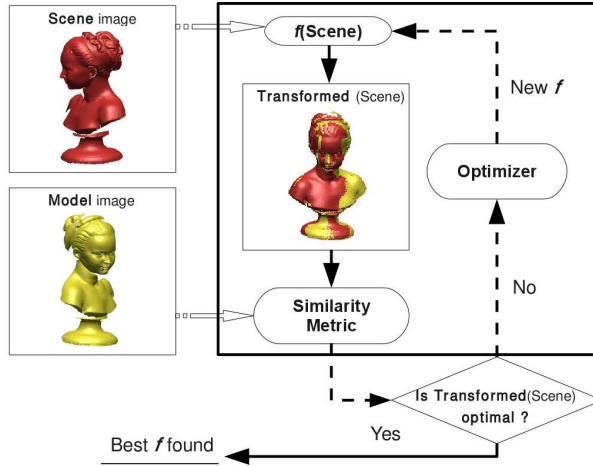
Figure 1: A general framework of the IR optimization process.

In this work, we will tackle two different IR domains: 3D model reconstruction and 3D MIR. Concerning the former, the 3D model reconstruction pipeline involves the application of several pair-wise alignments of two adjacent range images in order to obtain the final 3D model of the physical object [6]. In the last few years, specialized communities have experienced a growing interest using range scanners for building high-quality 3D models of real-world objects and scenes [6, 40] and avoiding humans to manually produce these models using laborious and error-prone CAD-based approaches [20]. Improved techniques within the 3D model reconstruction pipeline, e.g. IR algorithms, have been contributing to this task to date [12]. In this contribution, we adopt a two-stage pair-wise IR approach. It consists of a prealignment stage that provides coarse results, and a final refinement step, usually applying an ICP-based IR method, to slightly adjust the results. A rigid transformation which consists of a rotation and a translation of the images is usually considered for this environment.

Recently, the time-of-flight (ToF) technology is being used by novel real-time 3D video devices. Unlike traditional range scanners, ToF cameras have not being analysed enough tackling 3D model reconstructions. In this line of work, we proposed an image preprocessing pipeline to improve ToF image quality and analysed its performance applying an advanced BFOA design [5]. The same experimental setup will be considered in this work to test the proposed BFOA variants.

On the other hand, 3D medical IR is a mature research field where a wide variety of IR applications have been proposed. In medical applications, different 3D image modalities are usually acquired as magnetic resonance images (MRIs) and computer tomographies (CTs). In this framework, IR is a key technology that allows to "fuse" visual information from different sources [68]. Applications include combining images of the same subject from different modalities, detecting changes before/after treatment, aligning temporal sequences of images to compensate for motion between scans, image guidance during interventions, and aligning images from multiple subjects in cohort studies. There are excellent surveys that provide an up-to-date progress in the application of classical optimization techniques to the medical IR field [51].

Concerning the medical IR domain, we follow a single-stage IR approach, applying only one registration step. We will deal with complex scenarios by facing non-rigid IR problem instances considering similarity transformations, that include a rotation, a translation, and a uniform scaling. Such transformation has demonstrated its suitability in many medical image environments.

2.3. Evolutionary image registration

As stated before, the Optimizer component is of great importance in the success of the IR procedure. Since the ICP algorithm was introduced, many contributions have been proposed extending and partially solving its shortcomings (see Section 2.2) [30, 41, 55, 71]. Nevertheless, they still assume an initial near-optimal alignment of the images to be provided (the pair of images are usually pre-aligned manually before running any ICP variant).

EC [2] has demonstrated its ability to deal with complex real-world problems in CV and image processing. As an example, several issues on the topic have been published in international journals in the last few years [11, 47, 48]. The global optimization nature of evolutionary approaches, allows them to perform a roubst search in complex and ill-defined search spaces, demonstrating their ability to overcome traditional IR methods limitations, without requiring a good initial estimation of the image alignment.

In the last two decades, there is an increasing interest on applying EC principles to IR tasks due to their capability to provide global near-optimal solutions for complex optimization problems. The first attempts to face the IR problem using EAs can be found in the eighties [31], where a GA was developed for tackling rigid IR of 2D angiography images. Since then, evolutionary IR has become a very active area due to the successful results and several well-known EAs have been considered to tackle the IR optimization process in different domains and applications. For more information, we refer to [21], [57], and [65], which are three extensive reviews on this topic with 43 citations of evolutionary IR proposals including large experimental setups considering 13 different evolutionary IR methods.

There are different strengths and limitations of EAs when tackling complex optimization problems such as IR [21]. Some advantages are:

- Approaches based on EC and other MHs are conceptually simple and easy to implement. They are more robust than traditional gradient-based methods, as they do not depend on the starting solution. Moreover, they provide specific strategies to escape from local optima [70].

- It is easy to adapt the optimization scheduler of EAs to the particular constraint requirements of a wide range of possible objectives. They can also be integrated in a multi-objective scheme for solving the IR problem [62].

- EC and MHs have been used in a wide variety of optimization tasks within IR including numerical optimization and combinatorial optimization problems, i.e. facing both the transformation parameters and the matching-based IR approaches, respectively.

- The performance of the EAs is independent of the solution representation, unlike other numerical IR techniques that are only applicable for continuous functions or other constrained sets.

- It is easy to include prior knowledge about the problem to their design, yielding a more efficient exploitation of the geometry and the intensity information of the images [10, 17, 18].

- They can also be easily combined (e.g. hybridization) with traditional optimization techniques such as gradient-based methods [35, 59, 69].

The most important shortcomings related to these approaches are:

- Most EC approaches need a manual initial tuning of control parameters. Lately, some proposals include advanced strategies of adaptive control parameters to overcome this limitation [49].

- Typically, EC and MHs are time consuming. Parallel implementations have been proposed to face this pitfall [53].

- There is no formal proof to assure optimum convergence. Hybrid approaches are usually considered to escape from local optima [59].

- The estimation of the appropriate stop criterion is not easy and it is problem dependent. Either the CPU time or the number of function evaluations are typical criteria.

In some applications like 3D model reconstruction, a two-stage IR approach is usually considered (as seen in Section 2.2). First, a coarser and time consuming stage is faced using EAs, called pre-alignment. A second step, named refinement, is applied typically following an ICP-based IR algorithm[2]. Despite this, most evolutionary IR methods usually achieve accurate results at the pre-alignment stage (as we will see in the experimental section).

---

[2]Other refinement approaches have also been utilized [62].

3. BFOA-based Image Registration

This section is devoted to both introduce the different BFOA variants proposed and to describe their application to the IR task. To do so, we first present the common design decisions specifically associated to the application being solved and then elaborate on the proposed BFOA methods built to solve this problem. Notice that, these BFOA methods are generic and can be applied to different real-world applications.

3.1. Common Components: Coding scheme and objective function

Every pair-wise IR method involving the registration of two images aims to find the motion that brings the *scene* image (I_s) into the best possible alignment with the *model* image (I_m). The transformation considered will depend on the application addressed and the nature of the images involved. As we are tackling two different IR applications in this work to be solved using BFOAs, we will define two different optimization problems:

- As said, in RIR applications, an Euclidean motion is usually considered based on a 3D rigid transformation (f). This transformation is determined by six or seven real-coded parameters when using either Euler or axis plus angle representation for rotation, respectively [33]. Specifically, Euler-based rotation matrices suffer from the pitfall of the *gimbal-lock* and the specialized literature recommends using the second scheme [61] . Thus, we define the rigid transformation as: a rotation $R = (\theta, Axis_x, Axis_y, Axis_z)$ and a translation $\vec{t} = (t_x, t_y, t_z)$, with θ and $\vec{Axis}$ being the angle and axis of rotation, respectively. Then, the transformed points of the *scene* view are denoted by

$$f(\vec{p}_i) = R(\vec{p}_i) + \vec{t}, \quad i = \{1, \ldots, n\} \tag{5}$$

  where n is the number of points of the I_s image.

- In MIR applications, we face non-rigid IR problem instances by considering a 3D similarity transformation (f'), adding one more real-coded parameter to the representation, the scaling factor [33]. Thus, f' will be defined as a rotation R, a traslation $\vec{t}$, and a uniform scaling s, involving 8 real-coded parameters to be estimated. Then, transformed points of the *scene* view are denoted by

$$f'(\vec{p}_i) = sR(\vec{p}_i) + \vec{t}, \quad i = \{1, \ldots, n\} \tag{6}$$

The pair-wise IR procedure based on BFOAs can be formulated as an optimization problem developed to search for the euclidean transformation (f) in RIR applications, or the similarity transformation (f') in medical IR applications, achieving the best alignment of both $f(I_s)$ and I_m

$$f^* = \arg\min_f F(I_s, I_m; f) \quad s.t.: \ f^*(I_s) \cong I_m \tag{7}$$

according to the Similarity metric, F, being optimized. Any RIR solution is hence represented as a seven-dimensional (eight-dimensional for MIR) real-coded vector $x = \langle \theta, Axis_x, Axis_y, Axis_z, t_x, t_y, t_z \rangle$ (consequently, $x = \langle \theta, Axis_x, Axis_y, Axis_z, t_x, t_y, t_z, s \rangle$).

The median square error (MedSE) is usually considered as the F function in 3D model reconstruction [54] due to its robustness in presence of outliers (e.g. acquired noisy range images). It can be formulated as:

$$F(I_s, I_m; f) = MedSE(d_i^2), \quad \forall i = \{1, \ldots, n\} \tag{8}$$

where $MedSE()$ corresponds to the median value of all the squared Euclidean distances, d_i^2, between the transformed scene point, $f(\vec{p}_i)$, and its corresponding closest point, $\vec{p}'_j$, in the *model* view I_m, that is:

$$d_i^2 = \|f(\vec{p}_i) - \vec{p}'_j\|^2, \quad j = \{1, \ldots, m\} \tag{9}$$

where m is the number of points of the I_m image. Notice that, it can be said that both the F function and the fitness function have the same meaning within the optimization process. In order to speed up the computation of the closest point of every $f(\vec{p}_i)$ point, indexing structures as Kd-trees [71] or the grid closest point (GCP) transform [70] are often used. In this work we will use both indexing structures, the GCP structure to speed up the MedSE fitness computation and the Kd-tree structure to compare the final quality of the registration results.

7

3.2. Proposed BFOA variants

In this section several alternatives of improvement to be considered in BFOAs are introduced. All those BFOA variants are built by taking Dasgupta et al.'s proposal [23] as a base. They have been designed by fixing some specific decisions for some of the components and combining some other design choices for others, as described as follows:

- In the original proposal, Δ is a random vector which defines the direction of the tumble step within the range of each bacterium (see Eq. 1). In order to enhance the social behaviour of the bacteria simulated in the original swarming stage, and modify the tumble step negative effects, we propose an alternative solution in which a bacterium tends to follow the progress of another successful bacterium in the population, mimicking its direction.

- It may occur in the original reproduction step that the bacterium with less accumulated life (J_{health}) is not the best bacterium in the swarm. We propose an alternative based on a probabilistic selection of the best bacteria. Instead of sorting the bacteria in ascending order of their accumulated J_{health} values, a probabilistic filter is applied previous to sorting the bacteria. In first place, the current life value of each bacteria is normalized considering the life accumulated during the chemotaxis. Then, the best fit bacteria to be reproduced are selected with inverse probability of this normalized value. Finally, the bacteria are sorted and the first half reproduce as in the original method. This strategy increases the importance of those bacteria that obtain best fitness values in the last chemotactic cycles.

- Despite the elimination-dispersal step considers a small probability of application (1/8), we recognized the performance of the algorithm usually gets down, and it may lose some of the progress achieved in the chemotactic steps. Thus, we have considered an alternative using the best bacterium found so far to guide the dispersal step in order to minimize the negative impact of killing the best bacterium. Moreover, each transformation parameter has a 30% probability of being similar to the best bacterium.

- Besides, we introduce a set of elite solutions with the $M = 5$ best solutions found so far. After each chemotactic step, if the algorithm finds a solution, p_{new}, better than the worst elite solution, e_m, a recombination mechanism is applied considering p_{new} and a randomly selected elite solution. Recombination is done by using the $BLX - \alpha$ real-coding crossover operator [28]. In case the new solution, r_{new}, is better than p_{new}, the former will replace both the worst elite solution, e_m, and its parent. Otherwise, p_{new} will replace both e_m and its parent.

- Finally, we have used a local search (LS) strategy as a hybridization applied after the BFOA's reproduction step. In particular, we considered two different LS algorithms:

  - *Bound Optimization BY Quadratic Approximation (BOBYQA)*. It is an iterative algorithm based on the optimization of constrained functions using quadratic approximations [52]. These quadratic models are satisfied by interpolation conditions adjusted automatically by the algorithm. The interpolation points keep constant the separation distance, achieving precise local/global optima.
  - *Crossover Based Local Search*, known as XLS [8]. This LS is based on auto-adaptive crossover operators that fit the parent solutions of the population to generate new trial solutions. Given a parent BFOA solution, XLS randomly selects L bacteria of the population and it generates new ones in the parent neighbourhood using crossover operators.

3.3. Design of the proposed BFOA variants

In this section we present 52 different BFOA variants (Table 1) inspired in the design alternatives introduced in Section 3.2. Starting from the original proposals, A0 correspond to the Passino's canonical BFOA [50], A1 includes a dynamic evolution of the chemotactic step size to the previous method, and A2 correspond to the Dasgupta et al.'s adaptive BFOA variant [23]. The only difference between these three methods is calculation of the step size (see Eq. 4).

Then, we present four different configuration levels, integrating the previously proposed design alternatives, as follows: first level variants (A3-A6) use Dasgupta et al.'s BFOA as base scheme and integrate one of the proposed alternatives (i.e. the variation in the reproduction step, the alternative direction Δ, the elimination-dispersal step

variation, or the introduction of an elite set and the $BLX - \alpha$ operator, respectively). Second level variants (A7-A12) are combinations of two of the proposed alternatives. Then, third level variants (A13-A16) integrate three of the proposed alternatives. The fourth level variant (A17) consists on the combination of the four alternatives proposed in this work.

Finally, we extend the previous set of variants applying a hybridization to A1-A17 variants using two different LS strategies. Thus, A18-A34 are based on the BOBYQA local search, while A35-A51 make use of XLS. Notice that, when the LS procedure is considered, the BFOA variant becomes a memetic algorithm [46].

Algorithm	Description
Original BFOA methods	
A0	Passino's BFOA [50]
A1	A0 + dynamic evolution of the parameter C
A2	Dasgupta et al.'s BFOA [23]
First level variants	
A3	A2 + Reproduction variation
A4	A2 + Δ variation
A5	A2 + Elite set and $BLX - \alpha$ operator
A6	A2 + Dispersal variation
Second level variants	
A7	A3 + A4
A8	A3 + A5
A9	A3 + A6
A10	A4 + A5
A11	A4 + A6
A12	A5 + A6
Third level variants	
A13	A3 + A4 + A5
A14	A3 + A4 + A6
A15	A3 + A5 + A6
A16	A4 + A5 + A6
Fourth level variants	
A17	A3 + A4 + A5 + A6
Memetic variants	
A18	A1 + BOBYQA-based LS
A19	A2 + BOBYQA-based LS
A20-A34	A3-A17 + BOBYQA-based LS
A35	A1 + XLS-based LS
A36	A2 + XLS-based LS
A37-A51	A3-A17 + XLS-based LS

Table 1: Proposed BFOA variants (see Section 3.2).

4. BFOA Application to Range Image Registration for 3D Model Reconstruction

4.1. Range Image Registration Dataset

The range images used in the experiments were obtained from video sequences acquired using a PMD[Vision]® Camcube 2.0 time-of-flight (ToF) camera and three different objects were considered. Each image corresponds to the scan of objects at 45 degrees of rotation using a turn table. Then, every IR problem instance considers two range images of each object, i.e. facing a pair-wise RIR instance [5]. This is a really complex IR scenario due to the strong misalignment and the low overlapping between the two views. Figure 2 shows the snapshots of the three considered objects: Goblin, Teddy, and Pirate. The top row depicts the preprocessed range images.
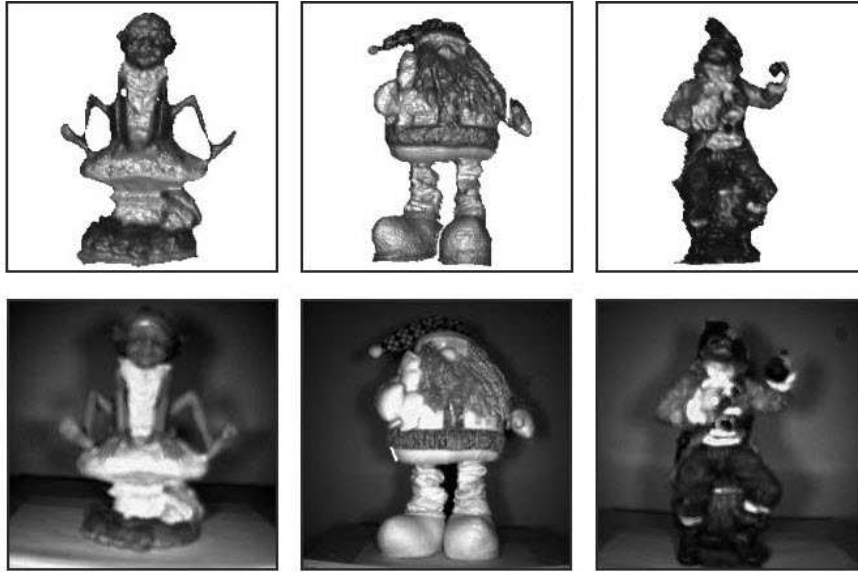
Figure 2: 3D/2D snapshots of the three scanned objects (*Goblin*, *Teddy*, and *Pirate*). Textured range images (left) and their corresponding photo (right) are shown.

Specifically, we applied three different image filters in order to enhance the quality of the raw images acquired by the PMD camera: noise reduction and the novel technique based on the *Shading Constraint* [9] principle. Next, a reduced subset of feature range points are extracted using specific extraction methods for 2D [4] and 3D [13] images. The IR method is run considering these subsets of points in order to achieve a fast convergence in the prealignment stage. Table 2 shows the number of points of both the preprocessed range image (raw data) and the extracted features using the third stage of the image processing pipeline.

RIR Dataset	*Goblin*		*Teddy*		*Pirate*	
	0°	45°	0°	45°	0°	45°
Raw data	11.468	10.983	13.802	13.641	11.340	10.210
Features	224	228	261	259	254	224

Table 2: Number of range points of the raw images and the extracted features for the RIR application.

4.2. Experimental Design

As said, each problem instance to be tackled consists of the registration of two adjacent views with strong misalignments (0° and 45°) for the three considered ToF images. We aim to test the performance of the BFOA variants described in Section 3.3. Besides, we considered the most relevant evolutionary and metaheuristic RIR methods [21] to set a high quality reference for the BFOA methods performance:

10

- Luck et al.'s [42] method, which is based on ICP and simulated annealing (ICP-L).

- Chow et al.'s [16] method. It makes use of real-coded GA (GA-C).

- Silva et al.'s [62] method, which makes use of an advanced steady-state GA design (GA-S).

- Wachowiak et al.'s [66] method. It is based on the particle swarm optimization (PSO) [36] algorithm (PSO-W).

- Santamaría et al.'s [58] method, which is based on the Scatter Search (SS) [39] algorithm (SS-S) and is currently the state-of-the-art evolutionary RIR method.

4.3. Parameter Settings

We considered the CPU time as the stop criterion for the prealignment stage (see Sections 2.2 and 2.3) in order to perform a fair comparison of all the IR methods. After a preliminary study, we determined that 20 seconds is a suitable value once noticed that every method properly converge to accurate enough solutions. In order to avoid execution dependence, 30 different runs of each IR algorithm have been performed for every problem instance. Additionally, we have set 40 iterations for the refinement stage which makes use of the ICP-based IR algorithm [71].

All the IR methods were run on an Intel® Core™ i7 2.00 GH platform with 10 GB RAM and implemented using C/C++ and the GNU/g++ compiler under the GNU/Linux Ubuntu 13.10 operating system. For every IR algorithm of the literature, we considered the parameter values recommended by the authors. Table 3 shows the parameter values used by the 52 proposed BFOA variants.

Parameter	Value	Description
n_p	7	Problem dimension.
n_s	16	Swarm size.
N_c	30	Number of chemotactic steps.
N_s	20	Number of swims.
N_{re}	6	Number of reproductive steps.
$Prob_{ed}$	0.125	Elimination-dispersal probability.
N_{ed}	3	Number of elimination-dispersal steps.
C_i	0.1	Step length (fixed only for Passino)
λ	0.0001	Constant value used to calculate the step length (C_i) in Dasgupta and HBFOA algorithms.

Table 3: Parameter settings of the BFOA variants.

4.4. Analysis of Results

This section presents the summarized results from which we analyse the performance of the proposed methods. This analysis is developed from two different viewpoints: the numerical optimization of the objective function and the empirical quality of the IR results, considering the prealignment (Section 4.4.1) and the refinement (Section 4.4.2) stages. The combined analysis of results provides us with more insights regarding to both the optimization and the specific problem solving branches for each considered dataset.

4.4.1. Analysis of the Prealignment stage

First, we carried out a broad study of the performance of the different configurations of BFOAs. As seen in Section 3.1, the considered BFOAs and state-of-the-art evolutionary IR algorithms tackle the prealignment of pairs of range images, and the numerical results are computed by using the MedSE metric as the corresponding fitness function. Specifically, we have chosen GCP [70] in order to speed up this first stage of the whole IR process. The MedSE values are used to compare the considered algorithms. Table 4 shows the outcomes obtained by all the 52 BFOA variants. Next, we analyse the results for each individual dataset.

According to *"Goblin"*, we found that:

- Concerning the comparison between the two basic BFOAS, A1 and A2, and the advanced variants, we found that 42 of 49 BFOA variants obtain better minimum results compared with the best minimum obtained by the canonical BFOA, A1. In addition, 38 of the 49 variants provide better mean performance compared with the best mean value obtained by the second standard BFOA, A2. Hence, we can see how the use of an advanced variant usually improves the standard BFOAs.

11

	Goblin			Teddy			Pirate		
	m	μ	σ	m	μ	σ	m	μ	σ
Standard BFOAs									
A0	0.04151	0.05036	0.00466	0.11081	0.14188	0.03814	0.03603	0.03886	0.00711
A1	0.04147	0.05000	0.00513	0.11171	0.12648	0.01871	0.03615	0.03737	0.00424
A2	0.04152	0.04838	0.00521	0.11154	0.11620	0.00166	0.03613	0.03667	0.00035
First level variants									
A3	0.04147	0.04491	0.00349	0.11340	0.11667	0.00119	0.03624	0.03679	0.00027
A4	0.04254	0.05035	0.00488	0.11429	0.11714	0.00116	0.03616	0.03690	0.00038
A5	0.04123	0.04369	0.00368	0.10980	0.11319	0.00121	0.03513	0.03604	0.00028
A6	0.04079	0.04632	0.00496	0.11342	0.11640	0.00106	0.03640	0.03677	0.00020
Second level variants									
A7	0.04153	0.04956	0.00513	0.11290	0.11680	0.00154	0.03603	0.03673	0.00036
A8	0.04123	0.04490	0.00392	0.11026	0.11286	0.00155	0.03574	0.03611	0.00010
A9	0.04121	0.04476	0.00350	0.11358	0.11675	0.00133	0.03608	0.03671	0.00025
A10	0.04140	0.04704	0.00419	0.11181	0.11477	0.00113	0.03553	0.03629	0.00026
A11	0.04233	0.05051	0.00482	0.11542	0.11740	0.00108	0.03617	0.03691	0.00035
A12	0.04089	0.04392	0.00394	0.10980	0.11287	0.00143	0.03584	0.03616	0.00025
Third level variants									
A13	0.04134	0.04731	0.00440	0.11152	0.11454	0.00097	0.03588	0.03632	0.00019
A14	0.04181	0.04919	0.00540	0.11081	0.11666	0.00179	0.03624	0.03691	0.00033
A15	0.04079	0.04331	0.00392	0.11026	0.11296	0.00124	0.03574	0.03605	0.00016
A16	0.04123	0.04690	0.00407	0.11217	0.11453	0.00094	0.03593	0.03634	0.00018
Fourth level variants									
A17	0.04133	0.04528	0.00400	0.11293	0.11469	0.00091	0.03590	0.03630	0.00020
Memetic variants using BOBYQA									
A18	0.04074	0.05106	0.00506	0.10952	0.14405	0.03923	0.03546	0.04392	0.01054
A19	0.04070	0.04983	0.00540	0.10726	0.11836	0.00861	0.03549	0.03675	0.00220
A20	0.04108	0.04822	0.00539	0.10779	0.11831	0.01016	**<u>0.03493</u>**	0.03663	0.00202
A21	**<u>0.04047</u>**	0.04621	0.00361	0.11045	0.11919	0.01030	0.03551	0.03612	0.00033
A22	0.04122	0.04803	0.00561	0.11065	0.11672	0.00640	0.03550	0.03627	0.00131
A23	0.04070	0.04905	0.00586	0.11052	0.11727	0.00498	0.03497	0.03619	0.00093
A24	0.04153	0.04956	0.00513	0.11290	0.11680	0.00154	0.03603	0.03673	0.00036
A25	0.04121	0.05019	0.00494	0.11110	0.12218	0.01008	0.03564	0.03689	0.00296
A26	0.04114	0.04867	0.00538	0.11026	0.11756	0.00978	0.03550	0.03687	0.00350
A27	0.04134	0.04818	0.00331	0.10926	0.11720	0.00668	0.03553	0.03617	0.00031
A28	0.04123	0.04673	0.00383	0.10853	0.11945	0.00977	0.03498	0.03613	0.00037
A29	0.04121	0.04744	0.00570	0.10905	0.11572	0.00505	0.03522	0.03625	0.00084
A30	0.04063	0.04678	0.00403	0.11066	0.11730	0.00616	0.03553	0.03617	0.00021
A31	0.04048	0.04793	0.00413	0.10916	0.11855	0.00815	0.03553	0.03620	0.00033
A32	**<u>0.04047</u>**	0.04915	0.00531	0.11171	0.11850	0.00700	0.03553	0.03631	0.00118
A33	0.04116	0.04710	0.00381	0.10933	0.11639	0.00671	0.03553	0.03611	0.00032
A34	0.04108	0.04697	0.00404	**<u>0.10720</u>**	0.11543	0.00451	0.03563	0.03618	0.00023
Memetic variants using XLS									
A35	0.04128	0.04552	0.00502	0.10980	0.12125	0.01178	0.03585	0.03615	0.00014
A36	0.04140	0.04510	0.00453	0.11181	0.11519	0.00125	0.03589	0.03646	0.00023
A37	0.04151	0.04485	0.00431	0.11260	0.11493	0.00124	0.03603	0.03637	0.00017
A38	0.04148	0.04561	0.00427	0.11171	0.11538	0.00117	0.03608	0.03657	0.00020
A39	0.04082	0.04377	0.00408	0.10980	0.11294	0.00129	0.03584	0.03611	0.00012
A40	0.04143	0.04431	0.00398	0.11338	0.11490	0.00096	0.03588	0.03637	0.00021
A41	0.04134	0.04585	0.00450	0.11354	0.11566	0.00094	0.03615	0.03648	0.00019
A42	0.04114	0.04567	0.00477	0.11007	0.11302	0.00106	0.03567	0.03605	0.00015
A43	0.04134	0.04417	0.00435	0.11081	0.11453	0.00132	0.03603	0.03631	0.00019
A44	0.04134	0.04497	0.00397	0.10980	0.11423	0.00167	0.03615	0.03635	0.00016
A45	0.04140	0.04540	0.00426	0.11221	0.11539	0.00142	0.03612	0.03657	0.00024
A46	0.04123	0.04387	0.00403	0.11157	0.11338	0.00081	0.03561	0.03607	0.00016
A47	0.04134	0.04409	0.00345	0.11217	0.11492	0.00095	0.03589	0.03634	0.00016
A48	0.04140	0.04524	0.00460	0.11354	0.11560	0.00101	0.03620	0.03658	0.00016
A49	0.04089	**<u>0.04328</u>**	0.00373	0.10952	**<u>0.11279</u>**	0.00141	0.03550	**<u>0.03600</u>**	0.00020
A50	0.04123	0.04463	0.00412	0.11081	0.11427	0.00109	0.03609	0.03637	0.00021
A51	0.04123	0.04476	0.00401	0.11340	0.11505	0.00089	0.03588	0.03636	0.00020

Table 4: Minimum (m), mean (μ) and standard deviation (σ) MedSE results for the RIR application. Results are scaled ($\times 10^3$).

- Regarding the comparison between the first to fourth-level variants, third level variants provide the better minimum and mean results, remarking the results obtained by A15.

- Comparing the memetic variants using LS and their corresponding first to fourth-level variant, 28 of the 34 hybrids obtain better minimum results, while 20 of 34 provide better mean performance.

- Concerning the use of the LS in the memetic variants, those using XLS obtain better mean performance with respect to those using BOBYQA. All the 17 hybrid proposals using BOBYQA but 3 obtain better minimum results.

- Overall, the best BFOA variants are memetics: A21 and A32 in minimum, and A49 in mean results.

Regarding to the *"Teddy"* dataset:

- Comparing the basic BFOAs and the advanced variants, 26 of the 49 variants achieve better minimum values compared with the best minimum obtained by the canonical BFOA, A0. Similarly, 26 of the 49 variants achieve a better mean performance compared with the best mean obtained by the second standard BFOA, A2.

- Regarding the comparison between the first to fourth-level variants, two different two level variants provide the better minimum (A12) and mean (A8) results.

- Concerning the memetic proposals using LS, we found that 24 of the 34 memetic variants obtain better minimum results than its corresponding first to fourth-level variant. In addition, 13 of the 34 memetic variants achieve a better mean performance.

- The memetic variants using XLS achieve better mean performance than those using BOBYQA. 14 of the 17 memetic variants using BOBYQA obtain better minimum values compared with those using XLS.

- Summarizing, the best BFOA variants make use of the memetic scheme. Specifically A34 and A49 obtain the best minimum and mean results, respectively.

The obtained results of the *"Pirate"* dataset, show us the following:

- Regarding the comparison between the basic BFOAS and the advanced variants, 37 of the 49 variants obtain better minimum values in comparison with the best minimum obtained by A0. Similarly, 37 of the 49 variants obtain better mean performance than the best standard BFOA, A2.

- Concerning the comparison between the first to fourth-level variants, the first level variants provide the better minimum and mean results, highlighting the results obtained by A5.

- Comparing the hybrid variants using LS and their corresponding first to fourth-level variant, 27 of the 34 memetic variants obtain better minimum results. In addition, 22 of the 34 memetic variants provide better mean performance.

- Concerning the use of the LS in the memetic variants, all the hybrids achieve similar minimum results. All the 17 hybrids using BOBYQA but 1 achieve better minimum results than those making use of XLS.

- Overall, the A20 and A49 memetic variants are the best variants and obtain the lowest minimum and mean values, respectively.

Next, we analyse the quality of the results using the Kd-tree [71] as closest point rule in both the prealignment and the refinement stages. Table 5 and Figure 3 show a comparison between the best BFOA variants and the state-of-the-art evolutionary and metaheuristic IR algorithms: ICP-L, GA-C, GA-S, PSO-W, and SS-S. The best performing BFOA methods are memetic: A20, corresponding to a first level variant based in A3 using BOBYQA as LS; A32, a third level variant based on A15 using BOBYQA; A34, a fourth level variant, A17, using BOBYQA; and A49, a third level variant based on A15 using XLS.

We can see how all the BFOA-based IR methods obtain better mean values than the baseline IR methods. Specifically, A34 and A49 variants obtain the best minimum and mean results in all cases but the *Pirate* instance, in which

PSO-W achieves the best minimum outcome. Notice that, this is a significant conclusion showing that the BFOA variants are a very good choice to design prealignment methods for the 3D model reconstruction pipeline (see Section 2.2). Figure 4 visually compares the registration results of the best IR methods according to the mean performance: A49 and SS-S. In order to present intuitively the visual results, both the model image (blue) and the transformed scene (yellow) have been coloured. An accurate registration will result in many coloured overlapped regions. Visual results during the prealignment stage show that both A49 and SS-S provide a precise registration for the three considered datasets.

	Goblin			Teddy			Pirate		
	m	μ	σ	m	μ	σ	m	μ	σ
A20	0.00579	0.01076	0.00307	0.01178	0.02367	0.01360	0.00655	0.00770	0.00109
A32	0.00569	0.01075	0.00284	0.01211	0.02034	0.01094	0.00597	0.00742	0.00070
A34	0.00584	0.01089	0.00280	**0.01162**	0.01607	0.00327	0.00606	**0.00725**	0.00041
A49	**0.00561**	**0.00760**	0.00275	0.01226	**0.01413**	0.00140	0.00647	0.00730	0.00052
ICP-L	0.01167	0.01314	0.00027	0.02788	14.7746	20.8366	0.00746	2.74628	0.50859
GA-C	2.49962	303.279	346.837	11.3260	440.990	498.899	3.29864	412.714	520.820
GA-S	0.00947	0.09527	0.07582	0.01638	0.29109	0.38431	0.00598	0.13806	0.37838
PSO-W	0.00842	0.01208	0.00087	0.01710	0.03689	0.01461	**0.00594**	0.00982	0.00276
SS-S	0.00568	0.00971	0.00340	0.01220	0.01446	0.00161	0.00653	0.00738	0.00039

Table 5: Minimum (m), mean (μ), and standard deviation (σ) MedSE values using Kd-tree in the prealignment stage. Results are scaled ($\times 10^3$).
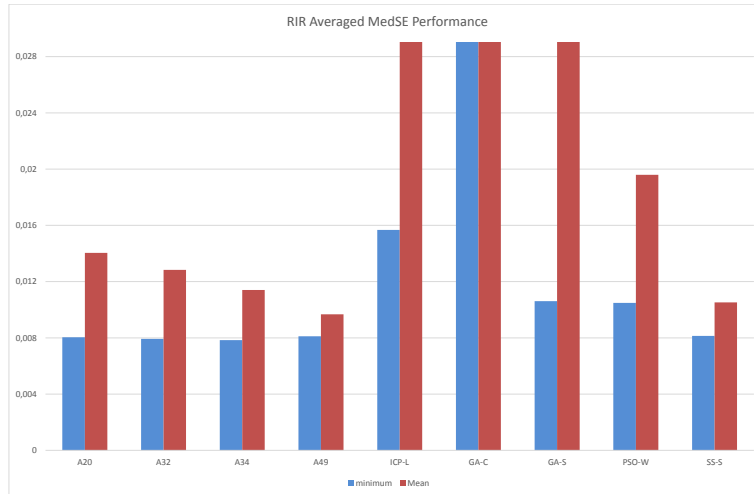

Figure 3: Averaged minimum and mean MedSE results using Kd-tree in the three datasets for both the best BFOA variants (A20, A32, A34, A49) and the state-of-the-art evolutionary and metaheuristic RIR algorithms in the prealignment stage. Results are scaled ($\times 10^3$).
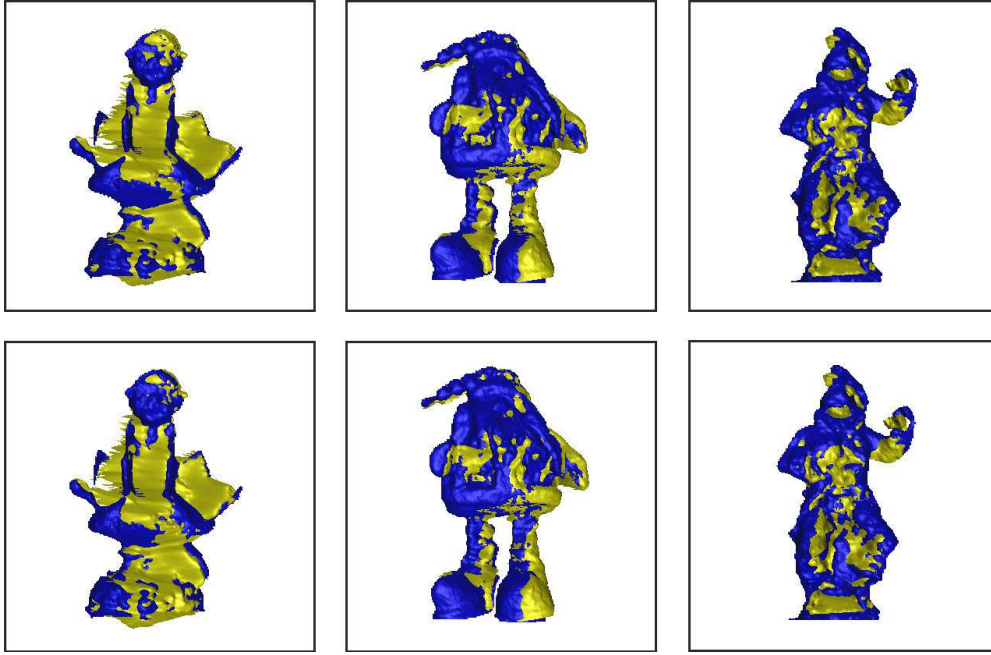
Figure 4: From left to right: First row corresponds to the best RIR estimations achieved by A49 tackling the *Goblin*, *Teddy*, and *Pirate* datasets in the prealignment stage. Second row shows the same results for SS-S.

4.4.2. Analysis of the Refinement quality

In this section, we extend the previous analysis by means of considering the second refinement stage to provide an accurate solution to the 3D model reconstruction problem (see Sections 2.2 and 2.3). A unified MedSE metric is considered to carry out a fair comparison between both stages. Table 6 shows the results after the refinement stage, in which it can be observed how the selected BFOA method achieve the best quality empirical results. All the RIR algorithms provide accurate enough prealignment solutions promoting the proper convergence of the refinement method. Specifically, A49 obtains the best mean performance in the three problem instances and the best minimum in the *Goblin* dataset, being the most robust IR method. On the other hand, SS-S matches the minimum result obtained by A49 in the *"Goblin"* dataset, while A22 and A32 achieve the best minimum in the *"Teddy"* and the *"Pirate"* datasets, respectively.

	Goblin			**Teddy**			**Pirate**		
	m	*μ*	*σ*	*m*	*μ*	*σ*	*m*	*μ*	*σ*
A20	0.00520	0.00850	0.00187	**0.01161**	0.01275	0.00167	0.00554	0.00579	0.00010
A32	0.00520	0.00802	0.00201	0.01164	0.01242	0.00132	**0.00545**	0.00578	0.00014
A34	0.00522	0.00844	0.00172	0.01163	0.01202	0.00015	0.00546	0.00575	0.00012
A49	**0.00518**	**0.00611**	0.00179	0.01169	**0.01198**	0.00015	0.00553	**0.00573**	0.00013
ICP-L	0.00934	0.00978	0.00008	0.01215	0.01284	0.00034	0.00579	0.00589	0.00002
GA-C	0.01296	0.07843	0.02832	0.02851	0.46021	0.23071	0.02499	0.07598	0.03782
GA-S	0.00525	0.06017	0.04673	0.01209	0.09637	0.08075	0.00547	0.03242	0.02784
PSO-W	0.00522	0.00942	0.00112	0.01199	0.01295	0.00144	0.00546	0.00586	0.00012
SS-S	**0.00518**	0.00700	0.00180	0.01172	0.01202	0.00015	0.00554	0.00577	0.00011

Table 6: Minimum (*m*), mean (*μ*), and standard deviation (*σ*) MedSE values using Kd-tree in the refinement stage. Results are scaled ($\times 10^3$).

These results demonstrate that the best BFOA variants (e.g. A49) provide a competitive performance compared to the state-of-the-art evolutionary and non-evolutionary IR methods. Figure 5 shows the averaged mean performance of every RIR algorithm in both the prealignment and the refinement stages. This graphic remarks the importance of obtaining an accurate enough prealignment result to promote a suitable convergence of the refinement process. Figure 6 shows the best refinement results for A49. The high quality reconstruction obtained can be clearly observed.
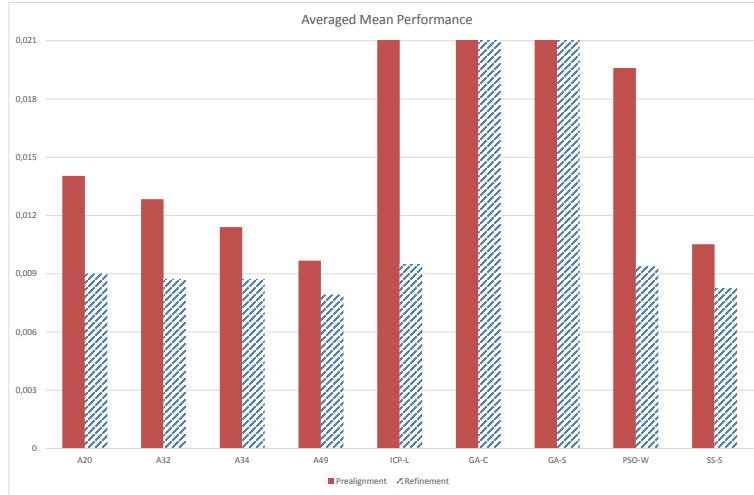
15

Figure 5: Averaged mean performance of the prealignment and refinement stages obtained in the three considered datasets for both the best BFOA variants and the state-of-the-art evolutionary and metaheuristic IR algorithms. Results are scaled ($\times 10^3$).
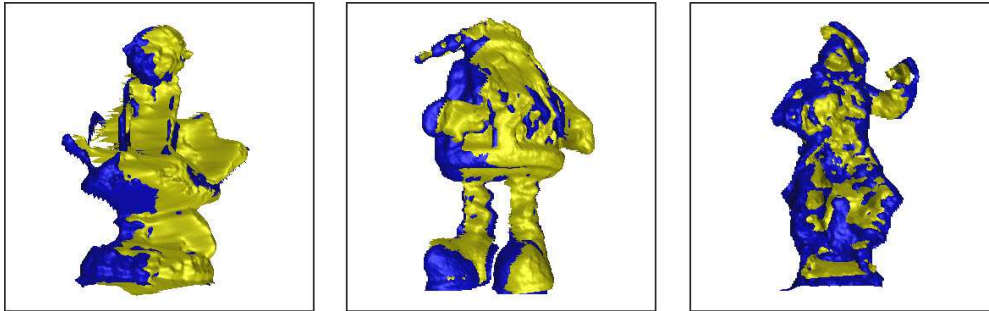

Figure 6: Refinement results of the best RIR estimations achieved using A49.

5. BFOA Application to 3D Medical Image Registration

5.1. Brainweb Medical Image Dataset

The medical images used in this work (see Figure 7) were extracted from the BrainWeb public repository [3] of the McConnell Brain Imaging Centre [38]. The BrainWeb repository provides synthetic MRI data computationally generated, based on two anatomical models: normal and multiple sclerosis (MS). Such MRIs have been extensively used by the neuroimaging community to evaluate the performance of different methods [3, 19, 63, 66].

For this dataset, a feature-based approach is considered using a 3D crest-line edge detector [45], which extracts the isosurface from points with relevant curvature information from the original images. Table 7 describes the particular settings considered to generate every medical image of our study (named MRIa, MRIb, and MRIc), detailing the nature of every medical image and the features extracted from each data set. In this contribution, each medical IR problem instance tackled considers two of the MRIs.

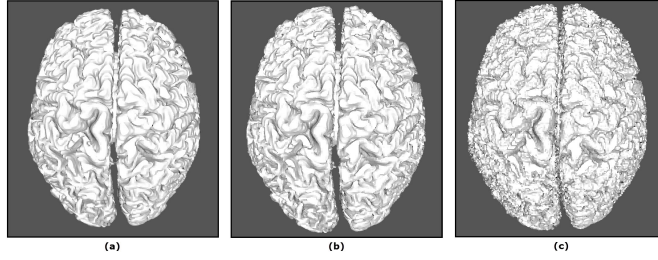[3] Available at http://www.bic.mni.mcgill.ca

Figure 7: MRIs extracted from the Brainweb image dataset. Noise levels correspond to 0% (a), 1% (b), and 5% (c) noise.

Medical IR Dataset	MRI model	Anatomical modality	level of noise	Raw data	Crest-line Features
MRIa	Normal	T1	0	12.644	584
MRIb	Normal	T1	1	11.493	393
MRIc	Lesion	T1	5	10.112	284

Table 7: Detailed description of the Brainweb image dataset and the extracted number of features.

5.2. Experimental Design

Each medical IR problem instance to be tackled consists of the registration of two medical images (MRIa-MRIb, MRIa-MRIc, and MRIb-MRIc), having three different scenarios overall. For this application, we analyse the behaviour of the 52 BFOA variants (Table 1) compared to the most relevant state-of-the-art evolutionary and meta-heuristic medical IR algorithms [21]:

- Luck et al.'s [42] method (ICP-L).

- He and Narayana's [34] method. It makes use of real-coded GA (GA-HE).

- De Falco et al.'s [29] method, which makes use of differential evolution [64] (DE-EA).

- Wachowiak et al.'s [66] method (PSO-W).

- Santamaría et al.'s [58] method (SS-S).

As for the previous experimental design, we considered the parameter settings originally proposed by the authors in every contribution and the same criteria shown in Section 4.3. In this case, the refinement stage is not considered as it is not usually applied for MIR applications [51]. To design a challenging scenario, a random (similarity) transformation is applied to the ground-truth image for each individual run of the algorithms and then the IR method estimates the unknown inverse transformation. Thus, a different ground-truth registration is available for each MIR problem instance in each of the 30 runs performed by all the IR methods. In particular, similarity transformations are randomly generated following a uniform probability distribution where each of the three rotation axis parameters will be in the interval $[-1, 1]$; the rotation angle will range in $[0, 360]$; the three translation parameters in $[-40mm, 40mm]$; and the uniform scaling in $[0.5, 2.0]$.

5.3. Analysis of Results

In this section we present the summarized results for the medical IR application. As in the previous experiments, we focus on both a numerical optimization and an empirical quality viewpoints during the analysis. First, we analyse the performance of all the 52 BFOAs (Table 8) considering the MedSE fitness function (see Section 3.1).

According to the first registration instance (*"MRIa-MRIb"*), we found that

17

	MRIa-MRIb			MRIa-MRIc			MRIb-MRIc		
	m	*μ*	*σ*	*m*	*μ*	*σ*	*m*	*μ*	*σ*
Standard BFOAs									
A0	0.35359	0.96110	0.11315	0.89954	0.98828	0.01717	0.60676	0.97820	0.06924
A1	0.93247	0.97055	0.01431	0.95821	0.98311	0.00914	0.93683	0.97735	0.01460
A2	0.40956	0.75297	0.22992	0.90275	0.95122	0.03107	0.66457	0.86532	0.12481
First level variants									
A3	0.43548	0.81875	0.20950	0.63204	0.89474	0.11866	0.64195	0.87781	0.11846
A4	0.36037	0.67959	0.24341	0.54638	0.88512	0.13887	0.59566	0.87053	0.13813
A5	0.39361	0.74419	0.25658	0.58497	0.84443	0.15888	0.63965	0.86478	0.13780
A6	0.39993	0.72424	0.23282	0.62091	0.84870	0.13670	0.65764	0.86448	0.12413
Second level variants									
A7	0.35583	0.73610	0.24750	0.55031	0.84668	0.15581	0.60677	0.90495	0.11219
A8	0.38322	0.72135	0.25957	0.60407	0.86959	0.14307	0.61315	0.86238	0.13371
A9	0.43005	0.75403	0.22592	0.63008	0.87153	0.13179	0.68728	0.87459	0.11365
A10	0.35236	0.73385	0.26326	0.54972	0.86692	0.15256	0.59564	0.83865	0.14958
A11	0.36037	0.74069	0.24450	0.53633	0.88894	0.13145	0.60677	0.85313	0.12823
A12	0.38320	0.71270	0.24702	0.58038	0.81666	0.16182	0.63293	0.84586	0.13468
Third level variants									
A13	0.35418	0.78072	0.25876	0.54635	0.86759	0.15321	0.60459	0.84270	0.14517
A14	0.35583	0.71343	0.23798	0.53488	0.82951	0.15360	0.60899	0.90143	0.11399
A15	0.37872	0.75703	0.25359	0.57295	0.85965	0.15172	0.62969	0.84040	0.13623
A16	0.35538	0.71391	0.26528	0.54972	0.83370	0.16633	0.59874	0.86340	0.13256
Fourth level variants									
A17	0.35263	0.76024	0.26690	0.54972	0.83795	0.17236	**0.59195**	0.85675	0.13726
Memetic variants using BOBYQA									
A18	0.93004	0.96911	0.01593	0.87068	0.97097	0.02414	0.91273	0.97410	0.01983
A19	0.35690	0.74659	0.27598	0.54769	0.88317	0.14915	0.62971	0.89071	0.13053
A20	0.35359	0.76122	0.26984	0.54974	0.88023	0.15148	0.59777	0.90572	0.12737
A21	**0.35060**	0.74191	0.27962	**0.52307**	0.86755	0.17151	0.59876	0.88588	0.14071
A22	0.35508	0.83107	0.22497	0.55028	0.87208	0.15593	0.61598	0.90093	0.12379
A23	0.35464	0.75767	0.26222	0.54769	0.85258	0.16112	0.62971	0.85896	0.13737
A24	0.35359	0.78871	0.25717	0.55031	0.89203	0.14821	0.59777	0.88448	0.14236
A25	0.35718	0.80946	0.24147	0.54635	0.87586	0.15011	0.60673	0.91479	0.11147
A26	0.35283	0.70441	0.29158	0.57412	0.89497	0.13463	0.59777	0.89193	0.13924
A27	0.35263	0.81539	0.24905	0.55024	0.84536	0.17411	0.59775	0.88198	0.14410
A28	0.35313	0.68957	0.28371	0.53488	0.87209	0.16229	0.59777	0.87041	0.15290
A29	0.35508	0.74254	0.27197	0.55028	0.86323	0.16642	0.60897	0.86867	0.13517
A30	0.35264	0.79222	0.25718	0.53457	0.84851	0.17675	**0.59195**	0.89423	0.13490
A31	0.35310	0.77980	0.26621	0.54976	0.87671	0.16042	0.60460	0.89760	0.12872
A32	0.35314	0.68794	0.28190	0.54635	0.84817	0.16615	0.60673	0.86822	0.13993
A33	0.35265	0.77931	0.26884	0.55024	0.87889	0.15789	**0.59195**	0.86767	0.15297
A34	0.35269	0.76000	0.27096	0.53355	0.87443	0.16380	**0.59195**	0.86972	0.14687
Memetic variants using XLS									
A35	0.92669	0.95708	0.01504	0.90384	0.95802	0.02273	0.92413	0.96134	0.01835
A36	0.43057	0.83073	0.20846	0.62592	0.86431	0.13846	0.67622	0.86515	0.11881
A37	0.43707	0.68013	0.23803	0.61143	0.87730	0.12928	0.66556	0.89187	0.11388
A38	0.35583	0.70046	0.24895	0.55031	**0.79885**	0.16414	0.61317	0.89790	0.11956
A39	0.38122	0.75872	0.24947	0.58309	0.85669	0.14927	0.65094	0.87910	0.12100
A40	0.42245	0.72494	0.24390	0.61963	0.83566	0.14476	0.66457	**0.82989**	0.12712
A41	0.35556	0.76778	0.24376	0.54638	0.90638	0.11835	0.60047	0.85598	0.13351
A42	0.40639	0.75936	0.24976	0.57295	0.83567	0.15482	0.64917	0.88958	0.12001
A43	0.43911	0.79556	0.22610	0.62783	0.85922	0.13650	0.66974	0.88419	0.11438
A44	0.35510	0.76677	0.25925	0.55028	0.85442	0.15672	0.59775	0.84234	0.15074
A45	0.36871	0.74065	0.25797	0.58203	0.84243	0.15286	0.59776	0.89576	0.12819
A46	0.35793	**0.61639**	0.23701	0.54491	0.82812	0.15644	0.63730	0.86563	0.12452
A47	0.35311	0.70565	0.26537	0.54635	0.85095	0.16829	0.59564	0.88697	0.12657
A48	0.35359	0.75268	0.24622	0.54975	0.88520	0.13655	0.60460	0.88349	0.12214
A49	0.39878	0.77357	0.24104	0.57916	0.84046	0.15807	0.63690	0.84284	0.13496
A50	0.35268	0.78070	0.25651	0.57457	0.86482	0.14877	0.59564	0.83619	0.14859
A51	0.35080	0.71419	0.27196	0.54972	0.85735	0.15690	0.59775	0.87601	0.13272

Table 8: Minimum (*m*), mean (*μ*) and standard deviation (*σ*) MedSE results for the Brainweb MRI dataset.

18

- Concerning the comparison between the standard BFOAs and the advanced variants, only 13 of the 49 variants achieve better minimum values compared with the best minimum obtained by the canonical BFOA, A0[4]. Similarly, 24 of the 49 variants achieve a better mean performance compared with the best mean obtained by the second standard BFOA, A2.

- Regarding the comparison between the first to fourth level variants, second and fourth level methods obtain better minimum results (A10 and A17, respectively), while a first level variant, A4, achieves better average performance.

- Comparing the memetic variants using LS and their corresponding first to fourth-level variant, 25 of the 34 hybrids obtain better minimum results, while 14 of 34 provide better mean performance.

- Concerning the use of the LS in the memetic variants, 13 of the 17 hybrid proposals using XLS achieve better mean performance than those using BOBYQA. In addition, 15 of the 17 hybrid variants using BOBYQA obtain better minimum values compared with those using XLS.

- Overall, the best variants make use of the memetic scheme. Specifically, A21 and A46 obtain the best minimum and mean results, respectively.

The obtained results of the *"MRIa-MRIc"* instance show:

- Comparing the basic BFOAs and the advanced variants, all BFOA variants (A3-A51) obtain better minimum results compared with the best minimum obtained by the canonical BFOA, A0. Similarly, 47 of the 49 variants obtain better mean performance compared with the best mean value obtained by the second standard BFOA, A2.

- Regarding the comparison between the first to fourth level variants, a third level variant (A14) provide the best minimum results. Furthermore, second level variants provide better mean results, highlighting the results obtained by A12.

- Concerning the memetic proposals using LS, 26 of the 34 hybrid variants obtain better minimum results than its counterpart first to fourth level variant. In addition, 19 of the 34 memetic variants achieve a better mean performance.

- The memetic variants using XLS obtain better mean performance in comparison with those using BOBYQA. All memetic proposals using BOBYQA but 2 obtain better minimum results in comparison with those variants using XLS.

- The best variants are memetic: A21 in minimum and A38 in mean results.

Regarding to the *"MRIb-MRIc"* instance:

- Regarding the comparison between the basic BFOAS and the advanced variants, 26 of the 49 variants obtain better minimum values in comparison with the best minimum obtained by A0. In addition, 18 of the 49 variants obtain better mean performance than the best standard BFOA, A2.

- Concerning the comparison between the first to fourth level variants, a fourth level variant (A17) obtains the best minimum result, while second level variants provide better mean results, remarking the results obtained by A10.

- Comparing the memetic variants using LS and their corresponding first to fourth-level variant, 19 of the 34 hybrids obtain better minimum results, while 10 of the 34 hybrids provide better mean performance.

---

[4]Notice that, the minimum values are a less significant indicator for this experiment as a different registration transformation is searched in each of the 30 runs performed by each IR method, as already mentioned.

- Concerning the use of the LS in the memetic variants, 14 of the 17 hybrids using XLS achieve better mean performance compared to those using BOBYQA. In addition, all the memetic variants using BOBYQA achieve better minimum results than those making use of XLS.

- Overall, A17, A30, A33 and A35 memetic variants obtain the lowest minimum values, while A40 is the memetic variant which achieves the best mean performance.

Table 9 shows a comparison between the best BFOA variants and the state-of-the-art evolutionary MIR algorithms: GA-HE, DE-EA, PSO-W, and SS-S. The best performing BFOA variants are all memetic: A21, corresponding to a first level variant based in A4 using BOBYQA as LS; A34, a fourth level variant, A17, using BOBYQA; A38, a first level variant based in A4 using XLS; A40, a first level variant based in A6 using XLS; and A46, a second level variant, A12, using XLS.

The best BFOA variants obtain competitive results compared to the best state-of-the art IR methods, DE-EA in average and SS-S in minimum results. Figure 8 depicts the averaged performance of the three considered instances regarding to the minimum and the mean results, respectively.

	MRIa-MRIb			*MRIa-MRIc*			*MRIb-MRIc*		
	m	μ	σ	m	μ	σ	m	μ	σ
A21	0.35060	0.74191	0.27962	0.52307	0.86755	0.17151	0.59876	0.88588	0.14071
A34	0.35269	0.76000	0.27096	0.53355	0.87443	0.16380	0.59195	0.86972	0.14687
A38	0.35583	0.70046	0.24895	0.55031	0.79885	0.16414	0.61317	0.89790	0.11956
A40	0.42245	0.72494	0.24390	0.61963	0.83566	0.14476	0.66457	0.82989	0.12712
A46	0.35793	0.61639	0.23701	0.54491	0.82812	0.15644	0.63730	0.86563	0.12452
GA-HE	0.67370	0.92009	0.07014	0.71123	0.93042	0.05905	0.81940	0.94485	0.03678
DE-EA	0.35038	**0.45045**	0.21427	0.52131	**0.68169**	0.19055	0.57298	**0.74460**	0.17087
PSO-W	0.98314	0.99540	0.00410	0.98765	0.99617	0.00359	0.99127	0.99729	0.00222
SS-S	**0.34941**	0.52681	0.27068	**0.52129**	0.70433	0.20847	**0.56538**	0.76076	0.18465

Table 9: Minimum (m), mean (μ), and standard deviation (σ) MedSE results of the best BFOA variants and the state-of-the-art evolutionary IR algorithms, considering the Brainweb MRI dataset.
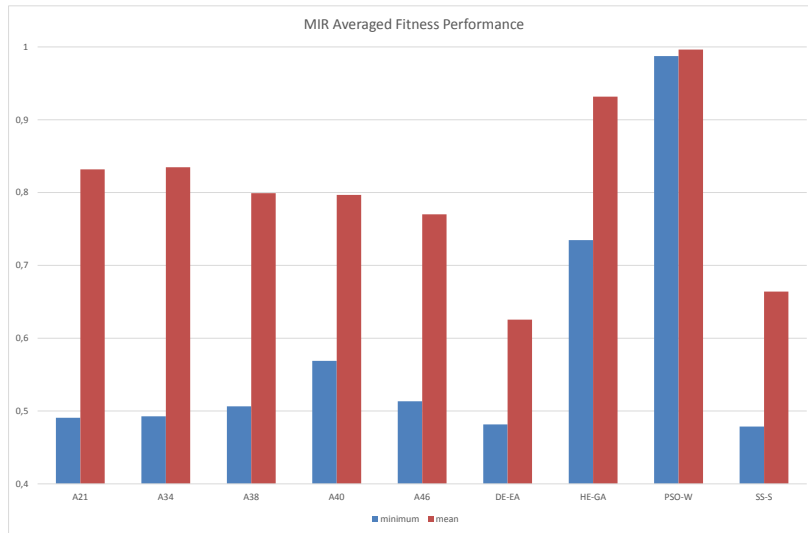

Figure 8: Averaged minimum and mean MedSE results of the three Brainweb MRI instances for both the best BFOA variants (A21, A34, A38, A40, A46) and the state-of-the-art evolutionary and metaheuristic IR algorithms.

20

Next, we analyse the final registration quality between the original images considering the unified MeanSE metric, more suitable for medical IR applications. This metric allows us to include ICP-L in the comparison, as it is not driven by the similarity metric used in the previous analysis. Table 10 shows the quality results of the best BFOAs compared to the state-of-the-art evolutionary and metaheuristic algorithms. In this case, BFOA variants A34 and A38 obtain the best minimum results in two instances, and DE-EA in the other one. Comparing the average performance, the results show how the best BFOA variants are usually the third best IR method, achieving competitive results with the most robust IR method, DE-EA, and outperforming ICP-L, GA-HE, and PSO-W. Figure 9 visually compares the results for the most robust BFOA variant, A46, and the best state-of-the-art IR algorithm, DE-EA.

	MRIa-MRIb			*MRIa-MRIc*			*MRIb-MRIc*		
	m	*μ*	*σ*	*m*	*μ*	*σ*	*m*	*μ*	*σ*
A21	7.52734	33.6204	30.5569	15.8257	40.9494	31.0017	14.5219	35.2496	29.6970
A34	7.54181	31.8119	31.6494	15.6649	40.7653	36.4613	**14.3074**	33.8253	30.6842
A38	7.58665	19.9195	19.5529	**15.6354**	27.1802	22.8515	14.7088	33.8536	24.3227
A40	7.73652	26.6422	36.9432	15.7340	34.8252	35.8491	14.3778	24.0626	20.5123
A46	7.58779	17.7286	19.8932	15.9884	27.4294	21.2792	14.3603	26.8315	20.2831
ICP-L	7.53203	75.2935	96.9296	15.7313	77.1457	64.7272	15.4881	38.5663	24.7864
GA-HE	10.9343	42.2374	53.9668	15.9136	37.8208	28.4723	14.7316	33.4822	38.7102
DE-EA	**7.52046**	**10.0073**	5.52872	15.8603	**19.0766**	5.06875	14.7249	**18.7715**	5.94312
PSO-W	115.685	2283.250	2215.23	70.3427	2353.43	2183.17	77.1475	2183.79	2130.71
SS-S	7.52120	14.64251	15.3293	15.9470	26.0639	21.3367	14.8053	20.9225	9.75530

Table 10: Minimum (*m*), mean (*μ*), and standard deviation (*σ*) MedSE values using Kd-tree considering the Brainweb MRI dataset.
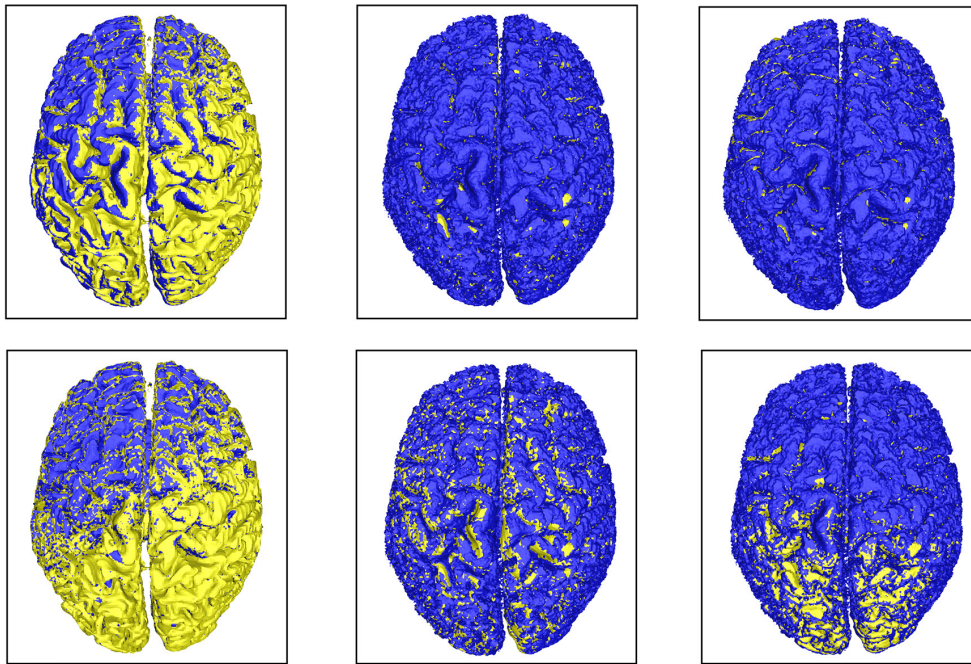

Figure 9: From left to right: First row corresponds to the best medical estimations achieved by A46 tackling the *MRIa-MRIb*, *MRIa-MRIc*, and *MRIb-MRIc* instances, respectively. Second row shows the same results for DE-EA.

These results demonstrate that the best BFOA variants (e.g. A46) provide a competitive and robust performance with respect to the considered state-of-the-art evolutionary and metaheuristic MIR methods, although at a significant distance of SS-S and DE-EA. Figure 10 shows the averaged mean performance of every algorithm over the three MIR scenarios. Notice that, all the BFOA variants outperform ICP-L, GA-HE, and PSO-W in every case, with A46 being the best performing one.
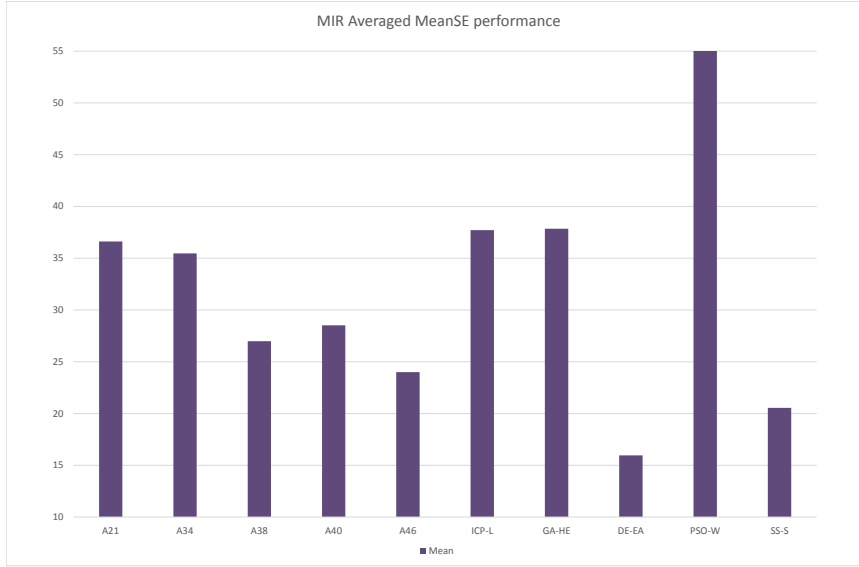
Figure 10: Averaged MeanSE results of the three Brainweb MRI instances for the best BFOA variants (A21, A34, A38, A40, A46) and the state-of-the-art evolutionary and metaheuristic IR algorithms.

6. General Discussion

In this section, we analyse the overall behaviour of the proposed BFOA variants for both IR applications. The first conclusion that can be drawn is that Dasgupta et al.'s BFOA [23] (A2) obtains an improved mean performance when compared to the other two standard BFOAs, A0 and A1. Thus, this clearly justify our variants (A3-A51) take A2 as their base scheme.

In the direct comparison between A0 and A1, Passino's BFOA [50] (A0) achieves better minimum results in every dataset but *"Goblin"*, where A1 achieves the minimum value.

Regarding the first level configurations of BFOA variants, the variation in the reproduction step, A3, only improves A2 in two of the six instances. The variation of the Δ operator developed in A4 has a negative impact in both minimum and mean results. On the opposite, the first level variant that makes use of elite set and the $BLX - \alpha$ crossover operator, A5, achieves the best minimum and mean performance within its category. Finally, the A6 variant, using the elimination-dispersal step, improves both minimum and mean performance in four of the six instances.

Considering the more advanced BFOA configurations from second to fourth levels, they outperform the canonical BFOA in the 75% of the cases, in both minimum and mean. In contrast, the advanced levels variants only improve their corresponding first level configurations results in the 30% of the cases. Hence, it is not only a matter of the simple addition of advanced components but a careful design is required.

Summarizing, the first to fourth level configurations show us that the best results are obtained by those using the $BLX - \alpha$ operator (i.e. A8, A12, and A15). Hence, this seems to be the most influential design decision in the advanced BFOAs. On the other hand, the negative effects of the Δ variation are propagated to more complex configurations: A7, A10, A11, A13, A14, A16, and A17. Besides, the variations in the reproduction and the elimination-dispersal steps do not clearly impact in the results.

According to the memetic variants, those using BOBYQA (A20-A34) obtain better minimum results in the 51% of the cases, and better mean performance in the 63%, in comparison to the memetic versions of the original BFOAs using BOBYQA (A18 and A19). Moreover, those memetic variants using XLS (A35-A51) obtain better minimum

results in the 61% of the cases, and better mean performance in the 62%, compared to the memetic versions of the original BFOAs using XLS (A35 and A36).

In general, 66% of the variants improve the best minimum results achieved by the original BFOA algorithms (A0 and A1), while 68% improve the best average performance achieved by the original BFOA (A2). The best minimum results for the six considered datasets are achieved by the memetic variants using BOBYQA, while the best average performance is provided by memetic variants using XLS (see Figure 11).
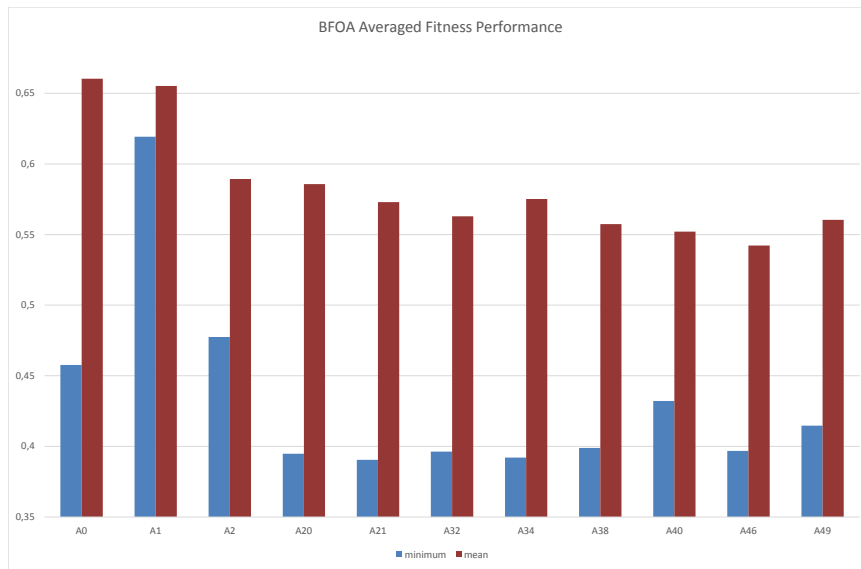
Figure 11: Averaged mean performance of the six considered IR problem instances for the best BFOA variants (A20, A21, A32, A34, A38, A40, and A49) and the standard BFOAs (A0, A1, A2). Results are normalized.

Among the best BFOA variants, A49 is the outstanding variant for range IR applications. It consists of a third level configuration (A15, based in A3, A5, and A6) using XLS, obtaining the best mean performance in the three considered RIR datasets. In the case of MIR applications, A46, a hybridization based on a second level configuration (A12, based in A5 and A6) using XLS, presents the most robust results.

From these results, its noticeable that the best BFOA configuration is different in each IR problem. In order to apply the different variants to other problem domains, it is necessary to study the specific design for each application, following the general rules presented in this work. Therefore, combining all the proposed variants into a memetic configuration using any LS is not enough to maximize the improvement in the performance. Taking that into account, memetic configurations using the $BLX - \alpha$ operator, the elimination-dispersal variation, and XLS are more likely to obtain a good performance.

Finally, we consider the quality of the registration for both applications, compared to the state-of-the-art IR algorithms. The results demonstrate that BFOA variants are a competitive approach when tackling both RIR and MIR problems, obtaining an accurate registration of the images. In some cases, they have been even able to outperform the best existing method.

7. Concluding remarks

In this work, our aim was to extend the state of the art of BFOA, a novel bioinspired algorithm based on the foraging behaviour of the E. Coli bacteria. Thus, we proposed a large amount of novel schemes resulting from the combination of different design decisions applied to the Dasgupta et al.'s self-adaptive BFOA version, obtaining improved variants.

Moreover, we have demonstrated the high performance achieved by these novel variants tackling IR, a complex real-world task in the computer vision field. Along the last two decades, evolutionary computation has demonstrated its capability to successfully tackle the IR problem and properly facing the drawbacks and limitations present in many of the classic IR methods. Hence, the choice of IR as an application field and of advanced evolutionary and metaheuristic IR methods as competitors results in a significant benchmark for BFOA design. In order to prove both the good performance and the suitability of the proposed BFOA variants from a general point of view, we considered two different domains within the IR problem, namely RIR for 3D model reconstruction and 3D medical IR.

From the results obtained in this experimental setup, we remark the high improvement obtained by some of the BFOA variants, which justify the election of BFOA as a valid swarm intelligence method to deal with real-world problems. In particular, many of the proposed variants obtained better results than those methods in the state-of-the-art, proving to be robust IR methods with really competitive results.

As future works, we aim to develop a comparative study of 3D reconstruction results using different acquisition devices based on the recent evolution of 3D camera devices. In addition, we aim to apply the designed BFOAs to some other complex optimization problems.

References

[1] D. Acharya, G. Panda, S. Mishra, Y.V.S. Lakshmi, Bacteria foraging based independent component analysis, in: International Conference on Computational Intelligence and Multimedia Applications, 2007, volume 2, pp. 527–531.

[2] T. Bäck, D.B. Fogel, Z. Michalewicz, Handbook of Evolutionary Computation, IOP Publishing Ltd and Oxford University Press, 1997.

[3] C. Baillard, P. Hellier, C. Barillot, Segmentation of brain 3D MR images using level sets and dense registration, Medical Image Analysis 5 (2001) 185–194.

[4] H. Bay, a. Ess, T. Tuytelaars, L. Vangool, Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding 110 (2008) 346–359.

[5] E. Bermejo, O. Cordón, S. Damas, J. Santamaría, Quality time-of-flight range imaging for feature-based registration using bacterial foraging, Applied Soft Computing 13 (2013) 3178–3189.

[6] F. Bernardini, H. Rushmeier, The 3D model acquisition pipeline, Computer Graphics Forum 21 (2002) 149–172.

[7] P.J. Besl, N.D. McKay, A method for registration of 3D shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (1992) 239–256.

[8] H.G. Beyer, K. Deb, On self-adaptive features in real-parameter evolutionary algorithms, IEEE Transactions on Evolutionary Computation 5 (2001) 250–270.

[9] M. Böhme, M. Haker, T. Martinetz, E. Barth, Shading constraint improves accuracy of time-of-flight measurements, Computer Vision and Image Understanding 114 (2010) 1329–1335.

[10] K. Brunnström, A. Stoddart, Genetic algorithms for free-form surface matching, in: International Conference of Pattern Recognition, Vienna, Germany, pp. 689–693.

[11] S. Cagnoni, E. Lutton, G. Olague, Editorial introduction to the special issue on evolutionary computer vision, Evolutionary Computation 16 (2008) 437–438.

[12] R.J. Campbell, P.J. Flynn, A survey of free-form object representation and recognition techniques, Computer Vision and Image Understanding 81 (2001) 166–210.

[13] U. Castellani, M. Cristani, S. Fantoni, V. Murino, Sparse points matching by combining 3D mesh saliency with statistical descriptors, Computer Graphics Forum 27 (2008) 643–652.

[14] H. Chen, Y. Zhu, K. Hu, L. Ma, Bacterial colony foraging algorithm: Combining chemotaxis, cell-to-cell communication, and self-adaptive strategy, Information Sciences In Press (2014).

[15] Y. Chen, G. Medioni, Object modelling by registration of multiple range images, Image and Vision Computing 10 (1992) 145–155.

[16] C.K. Chow, H.T. Tsui, T. Lee, Surface registration using a dynamic genetic algorithm, Pattern Recognition 37 (2004) 105–117.

[17] O. Cordón, S. Damas, Image Registration with Iterated Local Search, Journal of Heuristics 12 (2006) 73–94.

[18] O. Cordon, S. Damas, J. Santamaria, R. Marti, Scatter Search for the Point-Matching Problem in 3D Image Registration, INFORMS Journal on Computing 20 (2008) 55–68.

[19] P. Coupé, J. Manjón, E. Gedamu, D. Arnold, M. Robles, D. Collins, Robust Rician noise estimation for MR images, Medical Image Analysis 14 (2010) 483–493.

[20] G. Dalley, P. Flynn, Range image registration: A software platform and empirical evaluation, in: Third International Conference on 3-D Digital Imaging and Modeling (3DIM'01), pp. 246–253.

[21] S. Damas, O. Cordón, J. Santamaría, Medical Image Registration Using Evolutionary Computation: A Survey, IEEE Computational Intelligence Magazine 6 (2011) 26–42.

[22] S. Das, S. Dasgupta, A. Biswas, A. Abraham, A. Konar, On stability of the chemotactic dynamics in bacterial-foraging optimization algorithm, IEEE Transactions on Systems, Man, and Cybernetics, Part C 39 (2009) 670–679.

[23] S. Dasgupta, S. Das, A. Abraham, A. Biswas, Adaptive Computational Chemotaxis in Bacterial Foraging Optimization: An Analysis, IEEE Transactions on Evolutionary Computation 13 (2009) 919–941.

[24] S. Dasgupta, S. Das, A. Biswas, A. Abraham, Automatic circle detection on digital images with an adaptive bacterial foraging algorithm, Soft Computing 14 (2010) 1151–1164.

[25] A. Eiben, J. Smith, Introduction to Evolutionary Computation, Springer, Berlin, 2003.

[26] M. El-Abd, Performance assessment of foraging algorithms vs. evolutionary algorithms, Information Sciences 182 (2012) 243–263.

[27] A.P. Engelbrecht, Fundamentals of computational swarm intelligence, John Wiley & Sons, 2006.

[28] L.J. Eshelman, Real-coded genetic algorithms and interval schemata, in: L.D. Whitley (Ed.), Foundations of Genetic Algorithms 2, Morgan Kaufmann, San Mateo, EEUU, 1993, pp. 187–202.

[29] I.D. Falco, A.D. Cioppa, D. Maisto, E. Tarantino, Differential evolution as a viable tool for satellite image registration, Applied Soft Computing 8 (2008) 1453 – 1462.

[30] J. Feldmar, N. Ayache, Rigid, affine and locally affine registration of free-form surfaces, International Journal of Computer Vision 18 (1996) 99–119.

[31] J. Fitzpatrick, J. Grefenstette, D. Gucht, Image registration by genetic search, in: IEEE Southeast Conference, Louisville, EEUU, pp. 460–464.

[32] R.C. González, R.E. Woods, Digital Image Processing, Prentice-Hall, 2002.

[33] A. Goshtasby, 2D and 3D Image Registration, Wiley Interscience, 2005.

[34] R. He, P.A. Narayana, Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images, Computerized Medical Imaging and Graphics 26 (2002) 277–292.

[35] M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images, Medical Image Analysis 5 (2001) 143–156.

[36] J. Kennedy, R. Eberhart, Particle swarm optimization, in: IEEE International Conference on Neural Networks, volume 4, pp. 1942–1948.

[37] D.H. Kim, A. Abraham, J.H. Cho, A hybrid genetic algorithm and bacterial foraging approach for global optimization, Information Sciences 177 (2007) 3918–3937.

[38] R.K.S. Kwan, A.C. Evans, G.B. Pike, MRI simulation-based evaluation of image-processing and classification methods, IEEE Transactions on Medical Imaging 18 (1999) 1085–1097.

[39] M. Laguna, R. Martí, Scatter search: methodology and implementations in C, Kluwer Academic Publishers, Boston, 2003.

[40] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, D. Fulk, The digital michelangelo project: 3D scanning of large statues, in: ACM SIGGRAPH 2000, pp. 131–144.

[41] Y. Liu, Improving ICP with easy implementation for free form surface matching, Pattern Recognition 37 (2004) 211–226.

[42] J.P. Luck, C.Q. Little, W. Hoff, Registration of Range Data Using a Hybrid Simulated Annealing and Iterative Closest Point Algorithm., in: IEEE International Conference on Robotics and Automation (ICRA'00), pp. 3739–3744.

[43] T. Masuda, Registration and integration of multiple range images by matching signed distance fields for object shape modeling, Computer Vision and Image Understanding 87 (2002) 51–65.

[44] S. Mishra, C.N. Bhende, Bacterial foraging technique-based optimized active power filter for load compensation, IEEE Transactions on Power Delivery 22 (2007) 457–465.

[45] O. Monga, R. Deriche, G. Malandain, J.P. Cocquerez, Recursive filtering and edge tracking: two primary tools for 3D edge detection, Image and Vision Computing 9 (1991) 203–214.

[46] P. Moscato, On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms, Caltech concurrent computation program, Report 826 (1989).

[47] M. Nachtegael, E. Kerre, S. Damas, D. Van der Weken, Special issue on recent advances in soft computing in image processing, International Journal of Approximate Reasoning 50 (2009) 1–2.

[48] G. Olague, S. Cagnoni, E. Lutton, Introduction to the special issue on evolutionary computer vision and image understanding, Pattern Recognition Letters 27 (2006) 1161–1163.

[49] Y.S. Ong, M. Lim, N. Zhu, K. Wong, Classification of adaptive memetic algorithms: a comparative study, IEEE Transactions on Systems, Man, and Cybernetics, Part B 36 (2006) 141–152.

[50] K. Passino, Biomimicry of bacterial foraging for distributed optimization and control, Control Systems, IEEE 22 (2002) 52–67.

[51] J.P.W. Pluim, J. Maintz, M. Viergever, Mutual-information-based registration of medical images: a survey, IEEE Transactions on Medical Imaging 22 (2003) 986–1004.

[52] M. Powell, The BOBYQA algorithm for bound constrained optimization without derivatives, Technical Report NA2009/06, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 2009.

[53] C. Robertson, R.B. Fisher, Parallel Evolutionary Registration of Range Data, Computer Vision and Image Understanding 87 (2002) 39–50.

[54] M. Rodrigues, R. Fisher, Y. Liu, Special issue on registration and fusion of range images, Computer Vision and Image Understanding 87 (2002) 1–7.

[55] S. Rusinkiewicz, M. Levoy, Efficient variants of the ICP algorithm, in: Third International Conference on 3D Digital Imaging and Modeling (3DIM'01), Quebec, Canada, pp. 145–152.

[56] J. Salvi, C. Matabosch, D. Fofi, J. Forest, A review of recent range image registration methods with accuracy evaluation, Image and Vision Computing 25 (2007) 578–596.

[57] J. Santamaría, O. Cordón, S. Damas, A comparative study of state-of-the-art evolutionary image registration methods for 3D modeling, Computer Vision and Image Understanding 115 (2011) 1340–1354.

[58] J. Santamaría, O. Cordón, S. Damas, I. Alemán, M. Botella, A Scatter Search-based technique for pair-wise 3D range image registration in forensic anthropology, Soft Computing 11 (2007) 819–828.

[59] J. Santamaría, O. Cordón, S. Damas, J. García-Torres, A. Quirin, Performance evaluation of memetic approaches in 3D reconstruction of forensic objects, Soft Computing 13 (2009) 883–904.

[60] G.C. Sharp, S.W. Lee, D.K. Wehe, ICP registration using invariant features, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 90–102.

[61] K. Shoemake, Animating rotation with quaternion curves, in: ACM SIGGRAPH, San Francisco, pp. 245–254.

[62] L. Silva, O.R.P. Bellon, K.L. Boyer, Precision range image registration using a robust surface interpenetration measure and enhanced genetic algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2005) 762–776.

[63] D. Skerl, B. Likar, F. Pernus, A protocol for evaluation of similarity measures for rigid registration, IEEE Transactions on Medical Imaging 25 (2006) 779–791.

[64] R. Storn, Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces, Journal of Global Optimization 11 (1997) 341–359.

[65] A. Valsecchi, S. Damas, J. Santamara, Evolutionary intensity-based medical image registration: a review, Current Medical Imaging Reviews 9:4 (2013) 283–297.

[66] M.P. Wachowiak, R. Smolikova, Y. Zheng, J.M. Zurada, A.S. El-Maghraby, An approach to multimodal biomedical image registration utilizing particle swarm optimization, IEEE Transactions on Evolutionary Computation 8 (2004) 289–301.

[67] M. Wan, L. Li, J. Xiao, C. Wang, Y. Yang, Data clustering using bacterial foraging optimization, Journal of Intelligent Information Systems 38 (2012) 321–341.

[68] X.Y. Wang, S. Eberl, M. Fulham, S. Som, D.D. Feng, Data registration and fusion, in: D.D. Feng (Ed.), Biomedical information technology, Academic Press, 2008, pp. 187–210.

[69] X. Xu, R.D. Dony, Differential evolution with powell's direction set method in medical image registration, in: IEEE International Symposium on Biomedical Imaging: Macro to Nano, pp. 732–735.

[70] S.M. Yamany, M.N. Ahmed, A.A. Farag, A new genetic-based technique for matching 3D curves and surfaces, Pattern Recognition 32 (1999) 1817–1820.

[71] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, International Journal of Computer Vision 13 (1994) 119–152.

[72] B. Zitová, J. Flusser, Image registration methods: a survey, Image and Vision Computing 21 (2003) 977–1000.

3 Genetic algorithms for skull-face overlay including mandible articulation

- E. Bermejo, C Campomanes-Álvarez, A. Valsecchi, O. Ibáñez, S Damas, O. Cordón. Genetic algorithms for skull-face overlay including mandible articulation, Information Sciences, vol. 420, pp. 200-217, 2017. DOI: 10.1016/j.ins.2017.08.029

  - State: Published.
  - Impact Factor (JCR 2016): 4.832.
  - Category: COMPUTER SCIENCE, INFORMATION SYSTEMS. Order: 7/146. Q1.

Genetic algorithms for skull-face overlay including mandible articulation

Enrique Bermejo[a], C. Campomanes-Álvarez[a], Andrea Valsecchi[a], Oscar Ibáñez[a], S. Damas[b], Oscar Cordón[a,c,*]

[a]*Department of Computer Science and Artificial Intelligence, University of Granada, Granada 18071, Spain*
[b]*Department of Software Engineering, University of Granada, Granada 18071, Spain*
[c]*Research Center on Information and Communication Technologies, University of Granada, Granada 18071, Spain*

Abstract

Craniofacial superimposition (CFS) is a skeleton-based technique that aims to provide identity to a skull through its superimposition with one or more photographs of candidate missing people. While traditionally performed by forensic experts, computer-aided CFS methods can now provide substantial speedups and are quickly progressing towards a large degree of automation. A current major limitation concerns the position of the mandible, which is required to be manually set by the expert beforehand in order to reproduce the facial expression of the subject in each available photograph. This is time-consuming and prone to errors. In this work, we address this issue by extending the state-of-the-art genetic algorithm-based method with the ability to allocate the mandible in the right position according to an anatomical model. Based on a dataset of simulated ante-mortem images with different mandible apertures and facial poses, we prove experimentally that the proposed method is able to effectively tackle cases displaying a much larger range of mandible positions. In fact, thanks to the new genetic design, it is able to outperform the original method, even when the mandible aperture is very small.

Keywords: Forensic Identification, Craniofacial Superimposition, Skull-face Overlay, Mandible Articulation, Image Registration, Genetic Algorithms

1. Introduction

Craniofacial superimposition (CFS) [49] is considered one of the most relevant skeleton-based identification techniques within forensic anthropology [9]. It is a craniofacial identification [45, 42] approach in which a number of ante-mortem (AM) images of a missing person are superimposed over a skull to determine if they belong to the same subject by analyzing their morphological correspondence.

Three consecutive stages for the whole CFS process have been distinguished in [15]: 1) The acquisition and processing of the materials, i.e., the AM facial images and the skull (or 3D model of the skull). In some approaches, this step also involves the location of the craniometric and facial landmarks on the skull and the face, respectively. 2) The skull-face overlay (SFO), which focuses on achieving the best possible superimposition of the skull and a single AM image of the missing person. This process is iteratively repeated for each available AM image, obtaining different overlays. Skull-face overlay thus refers to what traditionally has been known as the adjustment of the skull size and its orientation with respect to the facial photograph [49, 29]. It is the most time consuming stage of the whole CFS procedure. 3) The decision making, where a degree of reliability for the craniofacial correspondence is established in order to show whether the skull and the available photograph belong to the same person or not (exclusion). This task requires a thorough analysis of the face/skull correspondence provided by SFO to determine if the skull and the face actually belong to the same person [14].

There is a strong interest in designing automatic methods to support the forensic anthropologist to put CFS into effect. In particular, the design of computer-aided CFS methods has experienced a boom over the past twenty years [24]. The most recent approaches use 3D models of the skull and soft computing (SC) methods for the first two CFS

---

stages [26, 27, 5] and include computer vision (CV) techniques to support the final decision [7, 8]. These methods allow us to both automate some tasks and handle their inherent uncertainty. In particular, the appropriate projection of the skull over the facial photograph is a very challenging and time-consuming part of the CFS technique. In fact, the estimation of the best correspondence for the SFO can take hours to reach the best possible fit [18, 43].

Our previous works tackle SFO automation using evolutionary algorithms and fuzzy sets [26, 27, 5]. These approaches are based on overlaying a 3D model of the skull over a facial photograph by minimizing the distance among pairs of landmarks while handling the imprecision introduced by the facial landmarks location [13, 4]. The minimization process involves the search for a specific projection of the skull model leading to the best possible matching between corresponding landmarks. In fact, reaching an optimal SFO accuracy is still an open field of research and a manual refinement is still required in many cases.

Up to now, all computer-based SFO methods have considered the mandible as a rigid (i.e., not articulated) part of the skull. Such a simplification causes a negative impact on the accuracy of the automatic SFO method. As the AM images used to perform CFS are typically provided by relatives, the missing person usually appears in relaxed situations, most of them smiling or with the mouth slightly open. Generally, cases with grimaces or forced poses are discarded due to the fact that the mandible is in an exaggerated position and these kinds of facial expressions distort the soft tissue of the face. Moreover, those photographs where subjects appear with their mouths open reduce the confidence of the identification. Therefore, it is essential to model the articulation of the mandible in order to improve CFS reliability, considering that, in most cases we only have skeletal information available to infer its movement. Moreover, the need to reproduce the position of the mandible as displayed in the AM photograph has been identified within the best practices in the CFS technical procedures [14] produced by the MEPROCS international working group[1]. Usually, a forensic expert proceeds to locate the mandible in order to approximate the facial expression of each one of the available photographs, either manually or using 3D modeling software.

As the capabilities of this novel approach are still unexplored, our intention is to extend the current SFO method [26, 27, 5] by incorporating a versatile and straightforward model of the mandible aperture movement. In this way, we simplify the parameterization of the problem, even at the expense of risking accuracy when representing the articulation displacement. Our approach can thus handle different mouth openings in the photographs and avoid the initial manual positioning of the skull (a tedious and imprecise task). We propose three variants of the genetic algorithm-based method to estimate the 3D/2D projection of the skull and its articulable mandible onto the available AM images. These three variants establish three different trade-offs between the search space size and the likelihood to achieve an optimal solution.

In order to analyze the performance of our methods, we consider an unbiased and reliable procedure based on the collection of a dataset of simulated AM images showing different apertures of the mandible. In particular, we consider 30 SFO problem instances from this ground-truth dataset, following a procedure similar to the one in [25]. To our knowledge, the current proposal is the first automatic SFO method that includes two major contributions: i) an estimation of the mandible aperture; and ii) a ground-truth dataset to validate the results.

The rest of this paper is organized as follows. In Section II we introduce the problem statement and the existing automatic SFO systems. In Section III we describe the articulated SFO algorithm proposal and detail its design to incorporate the aperture model. Section IV introduces the ground truth dataset we have designed to test and validate our proposal as well as the experimental setup, the analysis, and a discussion of the results obtained and their limitations. Finally, Section V draws some concluding remarks.

2. Preliminaries: State of the art in Automatic Skull-Face Overlay

From a general point of view, the SFO problem consists of positioning the skull in the same pose as the face in the photograph. In terms of CV, the AM photograph is the result of the 2D projection of a real (3D) scene that was acquired by a particular (unknown) camera [17]. In such a scene, the living person was in a given pose somewhere inside the camera's field of view. The most natural way to deal with the SFO problem is to replicate that original scenario. To do so, a 3D model of the skull must be used. The goal of this process is to adjust the size and the orientation of the skull

---

model with respect to the head in the photograph [6]. Besides, the specific characteristics of the camera must also be replicated to reproduce the original situation as much as possible. The latter is the approach followed by state-of-the-art automatic methods. In these approaches [26, 27, 5], the 3D skull model is positioned in the camera coordinate system through geometric transformations, i.e. translation, rotation and scaling, which corresponds to the adjustment of the skull size and its orientation at the same angle as the face in the image [49]. Finally, a perspective projection of the 3D skull model is performed onto the facial photograph. This is modeled as a 3D/2D image registration (IR) problem [26] whose solution is guided by a set of cranial and facial landmarks previously located by a forensic expert on both the 3D skull model and the facial photograph (see Fig. 1).
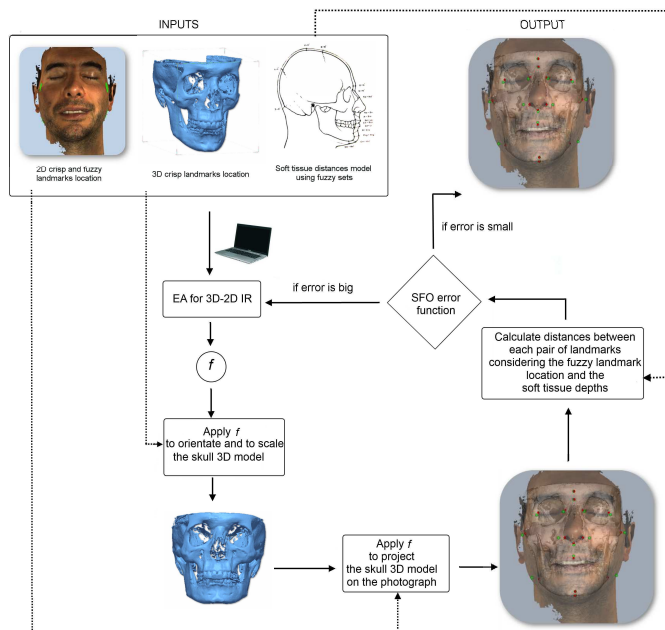

Figure 1: Scheme of the SFO procedure as a 3D-2D IR problem [6]

Once the location of these landmarks is provided by the forensic anthropologist, the SFO procedure is focused on searching for the skull orientation leading to the best matching of the two sets of landmarks. The required perspective transformation that needs to be applied to the skull was modelled in [26] as a set of geometric operations involving 12 parameters/unknowns encoded in a real-coded vector to represent a superimposition solution.

Hence, given two sets of cranial and facial landmarks, $C = \{cl^1, ..., cl^n\}$ and $F = \{fl^1, ..., fl^n\}$, the overlay procedure aims to solve a system of equations with the following 12 unknowns: the direction of the rotation axis $\vec{d} = (d_x, d_y, d_z)$, the location of the rotation axis with respect to the center of coordinates $\vec{r} = (r_x, r_y, r_z)$, the rotation angle θ, the factor s that scales the 3D skull model as the face in the photograph, the translation $\vec{t} = (t_x, t_y, t_z)$ that places the origin of the 3D skull model in front of the camera to replicate the same pose of the photograph, and the camera's field of view ϕ.

The authors employed different evolutionary algorithms to address the optimization process. Additionally, they extended the original objective function which minimizes the distances among cranial and facial landmarks to accommodate landmark location uncertainty (resulting from the difficulty to locate the facial landmarks in their exact locations using a photograph) [27] and landmark matching uncertainty (corresponding to the soft tissue distances existing between facial and cranial landmarks) [5] using fuzzy sets. As a result, the objective function to be minimized is defined as the Fuzzy Mean Error (FME):

$$FME = \frac{\sum_{i=1}^{Ncrisp} (d'(x_i, f(\tilde{C}^i)) + \sum_{j=1}^{Nfuzzy} (d''(\tilde{F}^j, f(\tilde{C}^j))}{N}, \qquad (1)$$

3

where *Ncrisp* is the number of 2D facial landmarks precisely located (crisp points), *Nfuzzy* is the number of 2D facial landmarks imprecisely located and defined as 2D fuzzy sets, N is the total number of landmarks considered ($N = Ncrisp + Nfuzzy$), x_i corresponds to a 2D facial landmark defined as a crisp point ($x_i \in F$), $\tilde{C}^i$ and $\tilde{C}^j$ are fuzzy sets modeling each 3D cranial landmark and the soft tissue distance to the corresponding 3D facial landmark i or j; f is the function that determines the 3D-2D perspective transformation that properly projects every 3D skull point onto the 2D photograph; $f(\tilde{C}^i)$ and $f(\tilde{C}^j)$ are two fuzzy sets, corresponding to the result of applying the perspective transformation f to the 3D volume, which models the landmark matching uncertainty; $\tilde{F}^j$ represents the fuzzy set of points of the imprecise 2D facial landmark; $d'(x_i, f(\tilde{C}^i))$ is the distance between a point and a fuzzy set of points, and $d''(\tilde{F}^j, f(\tilde{C}^j))$ is the distance between two fuzzy sets. Note that once estimated, f is applied to every 3D cranial landmark, handling the skull as a single rigid (i.e., not articulated) object.

3. Automatic Estimation of Mandible Aperture for Skull-face Overlay

In order to adequately model the mandible articulation in automatic SFO, it is essential to thoroughly analyze the mechanics of its movements. In Section 3.1, we review the existing mandible aperture models. Given the conditions to perform SFO, we select the best mandible aperture model for our aim in Section 3.2. Finally, we introduce the three proposed optimization method variants to estimate the transformations parameters encoding the SFO in Section 3.3.

3.1. Mandible aperture modeling

The Temporomandibular Joint (TMJ) connects the mandibular condyle to the temporal bone in the cranium, which are separated by a small oval fibrocartilage called articular disc [46]. This articulation is one of the most complex joints in the human body [39] as it features a bilateral articulation that can perform both gliding (translation) and hinge (rotation) movements [35, 36]. Both the articulation muscles and the articular disc allow the mandible to perform the different movements, while teeth occlusion, ligaments, and bones constrain the degrees of freedom of the joint [1] (see Fig. 2).
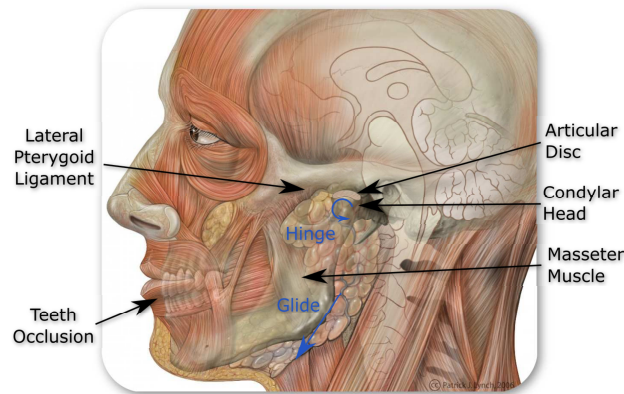

Figure 2: Illustration of the anatomy of the TMJ, by Patrick J. Lynch, medical illustrator, C. Carl Jaffe, MD, cardiologist. Used under Creative Commons Attribution 2.5 license 2006.

This complex movement has been widely studied in the oral biology and maxillofacial surgery fields. Most of the existing approaches try to represent the movement using biomechanical models [23, 47, 38, 32, 31], involving the stress forces induced by the muscles and the ligaments of the joint. Others attempt to explain this movement using a mathematical kinematic model, e.g. defining a kinematic axis [30], a kinematic center [48], an instantaneous center of rotation (ICR) [12], a screw axis [20], a helical axis [21], or a non-orthogonal floating axis [33]. Still, none of the latter proposals is able to faithfully simulate the anatomical joint, as all of them have limitations related to the geometric constructions and their calculations or inconsistences when determining the rotational axis, dealing with wide variations in the reproducibility of the opening movements [48, 11].

4

In Lemoine et al. [34] the authors used a 3D optical tracker to record mandibular movements during mouth opening and closing, and studied the relations of the mandibular geometry to define a prediction algorithm that can be applied to different case studies. They found a linear relationship between the length of the mandibular radius (i.e., the line between the center of the condylar head and the lower central incisor) and the mandibular movement. Thus, the mandibular radius is used as the scaling factor to fit the prediction model to each particular individual. Even though their experiments are based on a small sample of individuals, the results of their model concur with other studies based on recording in-vivo measurementes [3] or the proposal of a hierarchical ICR model to predict the articulation movement [44].

3.2. Selected mandible articulation model

In choosing the model, the main concern was the applicability to our SFO scenario. Even at the expense of accuracy loss, the selected mandible articulation model cannot be based on the interaction of the soft tissue, i.e. the position of the ligaments or the strength of the facial muscles. On the contrary, the movement of the mandible in SFO must be modeled using information that can be extracted from the bones only.

Lemoine et al.'s approach [34] offers a good compromise between reliability and simplicity when considering the parameterization of the model. In particular, mandible movement is modeled as a composition of: a) a rotation around the condylar axis (defined between the center of the two condylar heads) followed by b) a translation along the antero-posterior direction and c) a supero-inferior translation. The magnitude of the three movements depends on the aperture percentage (with 0% being teeth occlusion and 100% being maximum opening according to the normalized path shown in Fig. 3). It is important to highlight that up to 75% of the maximum opening may be approximated considering just the aforementioned rotation.
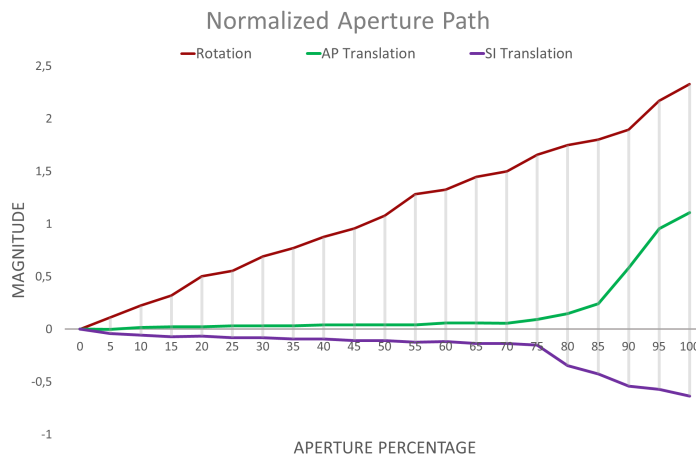

Figure 3: Aperture path graphic showing normalized rotation, antero-posterior and supero-inferior translation movements.

The magnitudes of the different movements are scaled using the mandibular radius, which can be easily identified in a radiographic analysis, measuring the distance from the center of the condylar head (condylar center) to the lower central incisor. In order to scale the normalized path, two linear regression equations are defined within the prediction model, relating the mandibular radius R_M with the maximum mandibular rotational angle (Θ_M) :

$$\Theta_M = (2.42)R_M - 1.24 \tag{2}$$

and with the maximum mandibular translation (T_M):

$$T_M = (-0.165)R_M + 2.91 \tag{3}$$

5

In our proposal, we obtain 3D models of the skull and the jaw for every identification case, either using a 3D scanner or a cone beam computed tomography (CBCT) scan. Then, an expert pinpoints the incisors and the mandibular and cranial condyles in the acquired 3D models during an initialization step. This approach avoids the use of a radiographic analysis for each case tackled. The process of determining this measurement in 3D is thus eased and the resulting error of locating the condylar center is reduced (Fig. 4).
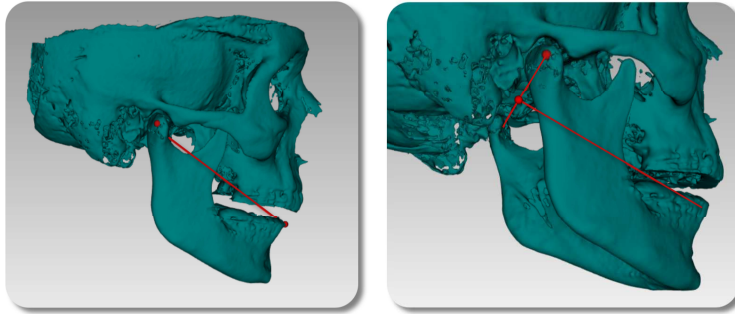

Figure 4: Illustration of the condylar center location and measurement of the mandible radius

The initialization step allows us to: 1) position the skull in a standardized anatomic pose (Frankfurt plane) [41], 2) automatically obtain the different magnitudes describing the mandibular opening movement for a given skull, and 3) directly apply the optimization method proposed in Section 3.3. By considering this step, the application of the technique is simplified by not imposing a manual positioning of the skull as in previous automatic and non-automatic SFO methods (see Section 4.5).

3.3. A Genetic Algorithm-based SFO method including mandible aperture estimation

Our proposal is based on the state-of-the-art genetic SFO method used in [26, 27, 5], including the objective function (Eq. 1). In order to incorporate the mandible articulation by means of the model described in Section 3.2, the mandible position is also controlled by an optimization parameter. This parameter simulates a different aperture percentage, where 0% corresponds to central occlusion and 100% to the maximum predicted opening in the sagittal plane. This method relocates the mandible landmarks prior to the application of the geometric transformation, adapting their position at every step of the optimization process, until the optimal solution is found. It is worth noticing that the complex SFO problem, which requires the estimation of 12 parameters, is only incremented by a single additional parameter concerning the mandible aperture, achieving a good trade-off between accuracy and optimization complexity.

The extension of the existing real-coded genetic algorithm (RCGA) for automatic SFO reviewed in Section 2 to include the mandible aperture parameter can be achieved in a number of ways. So, in order to look for the best design, we have tested three different variants, named A1, A2 and A3.

- **A1:** In this variant, the aperture percentage is directly added to the RCGA chromosome and thus undergoes the same evolutionary process as the 12 projection parameters (see Fig. 5). The optimization procedure ensures that the mandible position is estimated together with the skull transformation parameter in each candidate solution during each step of the genetic algorithm. A1 implements the most straightforward approach to extend the RCGA using an additional parameter. One of the advantages of this approach is that the full range of solutions of the extended problem is available. Nevertheless, this approach has an important drawback. The search space is increased and the good performance of the RGCA in optimizing the projection might be hindered by the interaction with the new parameter.
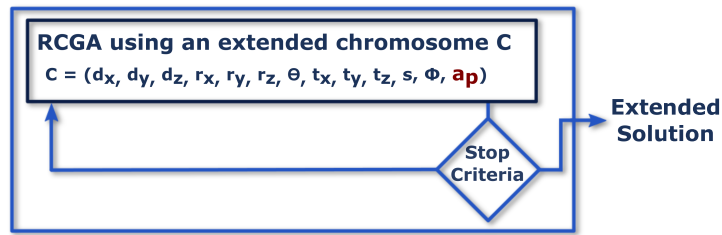
Figure 5: Coding scheme for A1 variant.

- **A2:** A local search to estimate the value of the aperture parameter is performed after the execution of the original RCGA method (see Fig. 6). In this case, the coding scheme and thus the search space size are kept unchanged for the global optimization procedure. Once a final solution is obtained, an exhaustive local search is applied so that all the possible mandible positions predicted by the aperture model are evaluated. It is intended as a refinement stage for the mandible position after the standard automatic SFO procedure, following a problem decomposition approach by means of a genetic local search. A2 is effective solely when the solution provided by the RCGA (i.e. not estimating the aperture) can be turned into a high-quality solution for the extended problem by adjusting only the mandible aperture. This is likely to happen when the SFO scenario has a small mandible aperture. The advantage of A2 over A1 is that more resources are devoted to optimizing the projection parameters. The inconvenience is that a partial solution is considered at each stage.
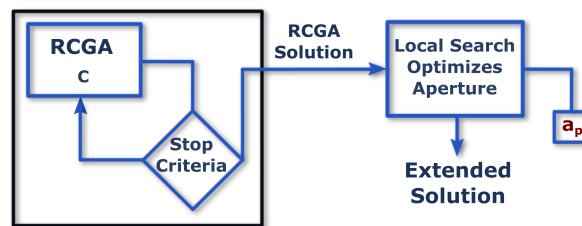

Figure 6: Coding scheme for A2 variant.

- **A3:** In this alternative, an exhaustive local search is applied after each generation of the RCGA in order to find the optimal aperture percentage parameter for the best individual found by RCGA, in a similar way to the A2 variant. The genetic evolution is applied to the original 12 unknowns only. Once a generation comes to an end, the optimal aperture for the best solution in the population is obtained and the individual is carried out to the next generation modifying its fitness values. A3 is based on a memetic algorithm considering an exhaustive local search applied over a single individual of the population. It thus achieves a compromise between the former two algorithms, by alternating the optimization of the projection and the mandible aperture, which could be more effective than doing it jointly.
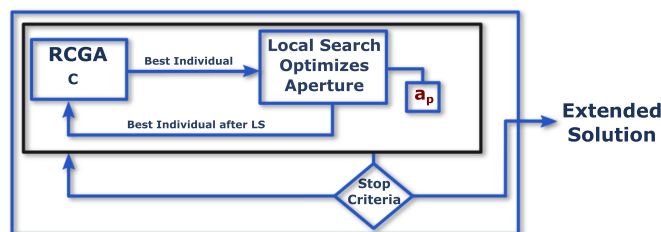

Figure 7: Coding scheme for A3 variant.

7

4. Experimentation

The objective of this experimental study is two-fold: i) to analyze the performance of the proposed articulated approach in comparison with the state-of-the-art automatic SFO method presented in [5]; and ii) to validate the obtained identification results, which is an essential requirement to verify the reliability and robustness of our new SFO approach.

Before presenting and analyzing the results of the experimental study, the procedure to create a ground truth dataset of SFO study cases is described. This dataset will allow us to analyze the performance of our methods, following an objective and reliable procedure.

4.1. Ground Truth Generation

Previous works in collaboration with forensic experts remark the importance of objective experimental analysis to validate the results and improve the confidence of the CFS technique [28]. In this proposal we aim to extend the ground truth dataset of [25] following a similar procedure. To that purpose, we generate a mandible aperture dataset with the collaboration of one subject, which was previously scanned using a CBCT with a neutral facial expression. The subject gave his consent to share his anonymized clinical data for research purposes.

In [25], 2D photographs for each individual were taken in a pose similar to the mandible pose during the CBCT acquisition, while the authors identified facial pose differences between the two acquisition times as a possible source of error. In this work, this issue is addressed by using simulated AM photos, projected from the actual 3D face model, which ensures a perfect correspondence with that 3D model. Although a single subject was involved, the generation of the data for the experimental study has been quite complex. The data of a single SFO case consist of a 3D model of the cranium and mandible (both segmented from the CBCT) together with an AM photo showing a specific mandible aperture and overall pose. The photo was simulated by rendering a 3D model of the face with a given mandible aperture in a specific pose (e.g. frontal). To this end, we acquired a series of 3D face models of the subject with different mouth apertures using a hand-held Artec Spider™ 3D scanner. Specifically, the subject was scanned with a 5 *mm* aperture variation (measuring inter-incisor distances). Ten scans were obtained ranging from central occlusion (rest) to the maximum aperture the subject was able to maintain comfortably during the scans (40 *mm*), including a scan of the subject smiling, to let the subject's teeth appear in the 3D model. A gauged resin cast was used to control the aperture distance for each scan (see Fig. 8).
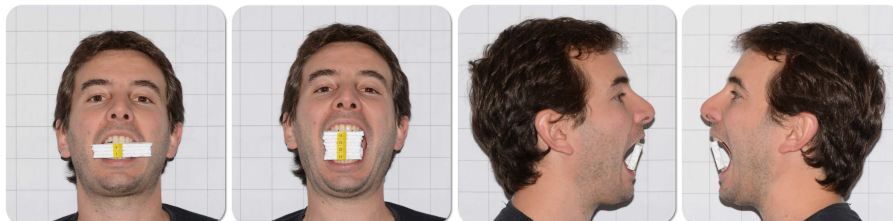

Figure 8: From left to right, the images show the detail of the resin cast used to control fixed aperture distances of 20 *mm*, 40 *mm*, 35 *mm*, and 40 *mm*, respectively.

The process described so far would be enough to create a dataset of SFO cases. However, we still need to compute the (a priori) solution of the cases, i.e., the projection f mapping the skull onto the photo. Note that we simulated the photo by setting the face model in a known pose (i.e., according to a transformation f') but this is not the same f that projects the skull. As the face and skull models are obtained from different sources, they are not aligned. To this end, we performed a manual segmentation of the CBCT to extract an additional face model (which is inherently aligned with the skull) and used it to align the dataset face models. Once the models are aligned, both f and f' are considered identical and provide the optimal solution used as a reference to measure the quality of the results. Finally, an expert marked the homologous cranial and facial 3D landmarks in the different models. Fig. 9 summarizes the data acquisition procedure for validation purposes.
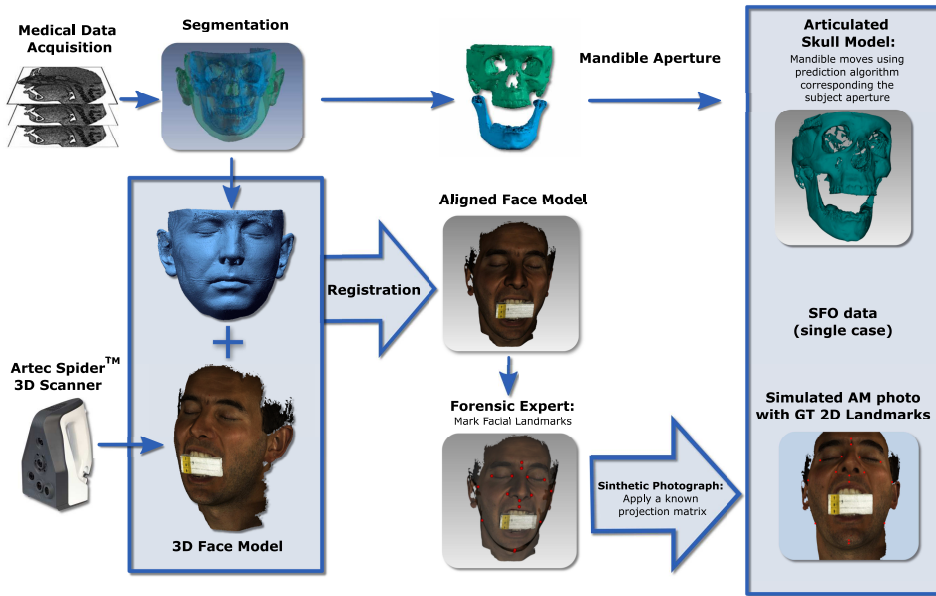
Figure 9: Overview of the Ground Truth data generation procedure

4.2. Experimental Setup

For each of the 10 facial scans in the ground truth dataset, three 2D photographs were extracted in different poses (frontal, lateral right, and lateral left) as projections of the 3D model. Thus, the experimental setup consists of 30 SFO instances that provide an objective testbed to analyze and evaluate the performance of our new approach in comparison with the state-of-the-art SFO automatic method, named RCGA [5].

All the experiments have been performed on an Intel Core^{TM} i7-3537U 2.0 GHz CPU, with 10 GB of RAM, running Windows 10 Professional^{TM}. For each SFO instance, 30 runs of each algorithm were considered, using a stopping criterion of 600.000 evaluations, and the same parameter configuration that yielded the best results in [5] (see Table 1). All the proposals that we consider make use of the same fitness function as the RCGA (see Eq. 1). Regarding the average execution time per run, there was a difference of 5 seconds between the fastest and the slowest algorithm: RCGA (25.05 s) and A1 (30.86 s), respectively.

Generations	600
Population size	1000
Crossover probability	0.9
Mutation probability	0.2
Tournament size	2
SBX η parameter	1

Table 1: RCGA parameter configuration.

4.3. Analysis of Experimental Results

In order to provide a fair and understandable comparison of the SFO results, we analyze two different error measurements. Once the skull is overlayed, we first calculate the error (mean Euclidean distance among landmarks) for all the estimated landmarks and their corresponding ground-truth positions. Thus, we compare the mean error of every SFO instance, which is obtained by averaging each landmark error, and the total average error for every instance of a particular algorithm. Furthermore, we calculate the inter-incisor distance between the upper and the lower central incisor in both the ground-truth 3D facial models and every SFO instance. This comparison allows us to determine the suitability and versatility of the mandible aperture model.

9

	Mean error				Aperture error		
Case, Pose	**A1**	**A2**	**A3**	**RCGA**	**A1**	**A2**	**A3**
Rest, F	1.54	1.60	**1.52**	1.55	0.27	0.07	**0.01**
Rest, L	3.23	**3.18**	3.22	3.24	**0**	**0**	**0**
Rest, R	3.20	3.11	3.16	**3.09**	**0**	**0**	**0**
Smile, F	**1.39**	1.40	1.52	1.52	0.81	**0.04**	0.45
Smile, L	**3.21**	3.28	3.34	3.26	**0**	**0**	**0**
Smile, R	**3.23**	4.33	4.23	4.12	**0**	0.66	0.37
5, F	**1.22**	1.30	1.52	1.30	4.87	1.93	**1.17**
5, L	3.69	2.89	**2.88**	3.71	**0.06**	3.89	3.74
5, R	**2.83**	2.93	2.92	2.90	**0.95**	3.86	3.60
10, F	**1.59**	1.72	2.16	1.86	3.21	**0.72**	2.14
10, L	**1.88**	2.44	2.32	3.71	**1.73**	6.67	6.53
10, R	3.48	2.74	**2.51**	2.99	**0,08**	6.70	5.93
15, F	**1.50**	1.96	2.61	2.54	4.94	1.23	**0.24**
15, L	**2.44**	3.04	3.14	3.96	**2.09**	7.82	8.82
15, R	**2.58**	3.11	3.38	4.02	**2.28**	7.64	9.03
20, F	**1.30**	2.40	2.80	3.27	**2.83**	8.99	6.08
20, L	**2.91**	3.92	4.02	4.93	**1.57**	11.09	10.28
20, R	**3.11**	3.65	3.92	4.95	**1.84**	9.31	10.87
25, F	**1.64**	2.45	3.25	3.47	3.93	6.92	**3.01**
25, L	**3.74**	4.66	5.01	5.66	**0.83**	12.41	15.73
25, R	**2.58**	4.79	5.51	5.79	**2.59**	11.89	12.41
30, F	**1.28**	2.98	3.76	3.84	**2.40**	15.07	7.43
30, L	**2.51**	4.22	5.69	6.56	**0.73**	13.19	13.38
30, R	**2.55**	3.83	5.87	6.60	**0.49**	10.95	14.19
35, F	**1.24**	3.10	3.74	4.04	**0.97**	19.69	9.98
35, L	**3.71**	4.52	6.23	7.35	**1.88**	16.91	16.66
35, R	**2.40**	4.66	5.91	7.44	**2.31**	17.02	15.72
40, F	**1.67**	3.53	4.68	5.01	**0.88**	18.92	13.86
40, L	**4.57**	6.75	6.90	8.29	**0.28**	18.74	22.49
40, R	**2.38**	8.14	7.80	9.88	**1.81**	17.70	19.32

Table 2: Mean and aperture distance error for every SFO instance of the ground truth dataset.

Table 2 presents the mean error per algorithm and the average inter-incisor distance error for each instance, distinguished by case (actual incisor aperture value) and pose (F= frontal, R = right, and L= left views), for the three proposed algorithms considering mandible articulation and the original RCGA. Both *Rest* and *Smile* cases correspond to an aperture of approximately 0 *mm* (closed mouth in central occlusion). The only difference is that, in the latter, the subject is smiling, so his teeth are visible in the 3D model taken as reference.

Regarding the mean error (left-hand side of Table 2), the results show that A1 stands out as the best performing algorithm considering the ground truth in all but 5 instances, while the reference algorithm, RCGA, only outperforms the proposed variants in 1 of the 30 instances. The mean error of the best solutions ranges from 1.22 to 4.57 *mm*. Considering only the three proposed variants, A3 obtains the best mean error in 3 of the 30 instances, while A2 only achieves the best value in one occasion.

It is worth noticing that the error values obtained are always lower in the frontal instances (< 1.7 *mm*), while lateral poses tend to increase the resulting error (> 1.8 *mm*). This behavior is explained by the inherent complexity of the lateral instances, as there are less visible landmarks in the 2D photograph, hindering the optimization process (readers interested in coplanarity problems in image registration are referred to [40] for a deeper explanation). Notice also that when the aperture is small (i.e., 0 − 10 *mm* aperture range), the differences among the results are negligible. This behavior is explained by the fact that the local search variants (A2 and A3) only focus on the aperture percentage parameter. They thus improve the mandible position obtained by the basic algorithm scheme with lower apertures. Nevertheless, they are unable to properly determine the skull location when the aperture is larger. That is the case

of all the remaining problem instances, where A1 reduces the resulting error notably (more than 0.45 *mm*). This improvement confirms our hypothesis that A1 would be able to reach more accurate solutions since it explores the whole search space at the expense of increasing the optimization complexity. In view of these results, the inclusion of the aperture percentage parameter within the evolutionary optimization process is strongly recommended.

Fig. 10 provides an overview of the results in terms of mean error. Note that the three plots represent the same data from three different viewpoints. Overall, the range of error values for cases having lateral poses is much larger than that of frontal ones. This is also reflected in the corresponding mean values. However, most solutions are actually quite close to the mean, with a few solutions scoring very high errors. In terms of variability, A1 shows the most consistent behavior. In comparison, A2, A3, and RCGA delivered a broader range of results, with A2 being the worst in the group.

Considering the averaged inter-incisor distance error, the right-hand side of Table 2 does not include RCGA results because mandible aperture is fixed at the original aperture extracted from the CBCT. By looking at the inter-incisor distances (Fig. 11), one can draw the very same conclusions as for the mean error. The superiority of A1 is even more clear in this occasion. These results highlight the overall best performance of the A1 proposal, especially when considering larger apertures. A1 achieves a better performance in lateral views, where the resulting aperture error is always lower than 2.6 *mm*. On the contrary, A1 aperture error is up to 5 *mm* in some frontal views with lower aperture.

In order to complement the analysis of the results, we applied different post-hoc statistical analysis [22], which are summarized in Table 3. First, we considered Friedman's nonparametric test [19], which aims to test the null hypothesis stating that there is no difference among the resulting mean errors of all the algorithms. The result of applying Friedman's test is $\chi^2_F = 44, 52$, and the corresponding p-value is $1, 17 \cdot 10^{-09}$. Since the p-value is much lower than the considered level of significance ($\alpha = 0.05$), the test concludes that there are significant differences among the observed results.

Algorithm	Mean Error	Ranking	Bonferroni-Dunn p	Holm p
A1	2.49	1.36	-	-
A2	3.42	2.26	0,02	0,02
A3	3.85	2.86	$< 10^{-4}$	$< 10^{-4}$
RCGA	4.36	3.50	$< 10^{-9}$	$< 10^{-9}$

Table 3: From left to right: Mean error of each optimization algorithm, ranking obtained through Friedman's test, statistical p-values on the results with A1 as control method for Bonferroni-Dunn's test, and for Holm's test.

Next, we considered the Bonferroni-Dunn test [16] to detect significant differences between a control approach and the rest of the algorithms. In this case, the control algorithm is the proposed variant A1 and the results are graphically represented in Fig. 12. A Bonferroni-Dunn's graphic illustrates the significance of the difference among rankings obtained for each algorithm considering the mean error. The horizontal line shows the significance threshold at $\alpha = 0.05$, whose height corresponds to the sum of the ranking of the control method (A1) and the critical difference value computed by the Bonferroni-Dunn method. Thus, algorithms exceeding the threshold line perform significantly worse than the control method for the considered α.

The Bonferroni-Dunn's test shows us that the control method A1 is significantly better than every other algorithm at $\alpha = 0.05$ with an ample performance gap. Moreover, we complemented the statistical analysis with another powerful test, such as Holm's, while still comparing A1 with the rest of the methods. The p-value results for Bonferroni-Dunn's and Holm's procedures are included in Table 3, providing similar results with p-values smaller than $\alpha = 0.05$.

4.4. Qualitative Analysis of Results

A visual assessment of the quality of the results is depicted in Figs. 13 and 14 (including frontal and lateral cases, respectively). Both figures facilitate a comparative visual analysis between the best SFO solution (i.e., the solution achieved by A1 with the lowest mean error in the 30 runs) and the solution obtained by the state-of-the-art RCGA method. Each visual SFO solution consists of the superimposition of the skull and jaw 3D models over the simulated photograph. It also represents the resulting projections of two set of points: cranial landmarks (in green circles) and facial landmarks (in red boxes).
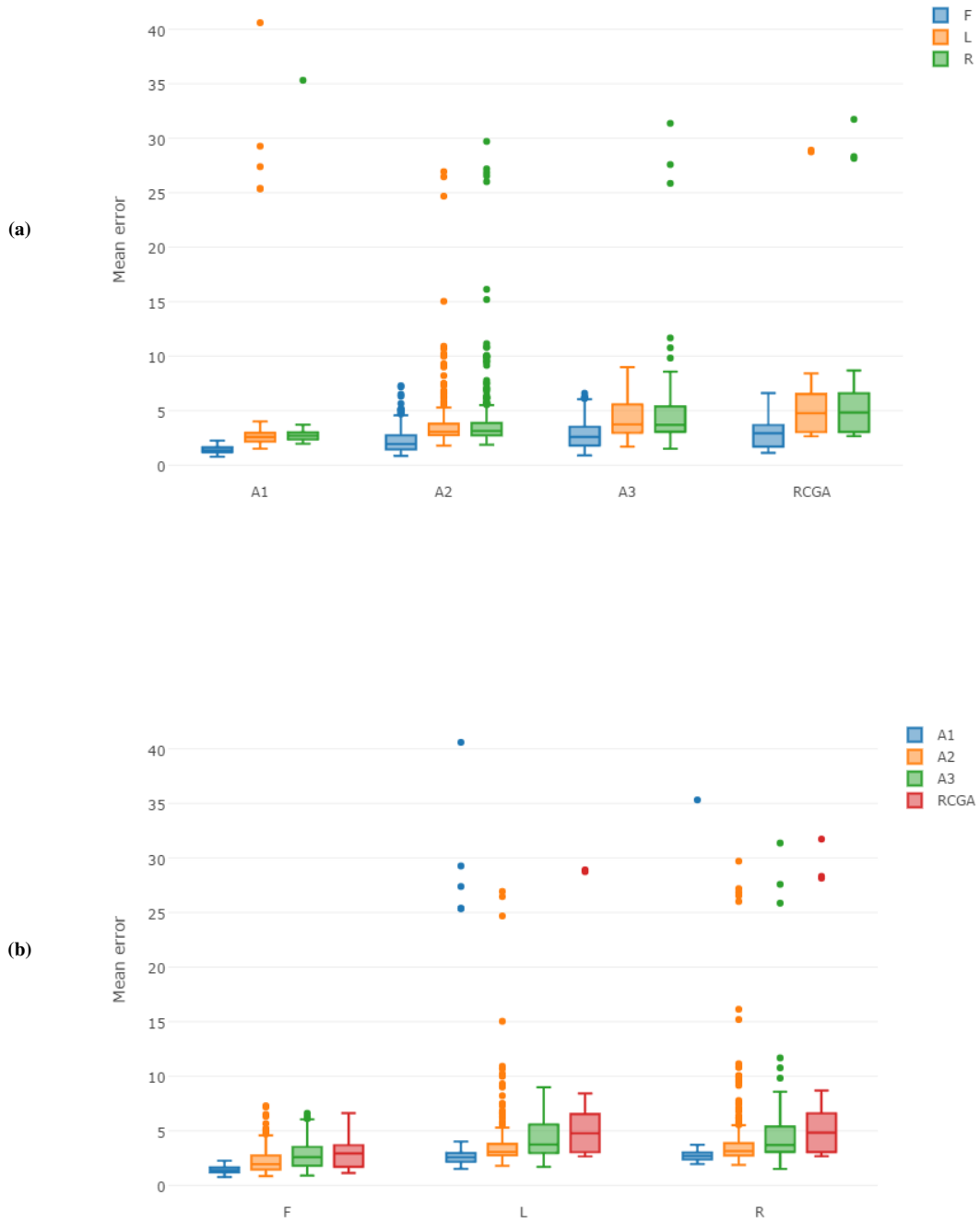
Figure 10: Box-plots corresponding to mean error (in mm) grouped by : (a) algorithm and pose, (b) pose and algorithm, (c) algorithm and mandible aperture.

(c)

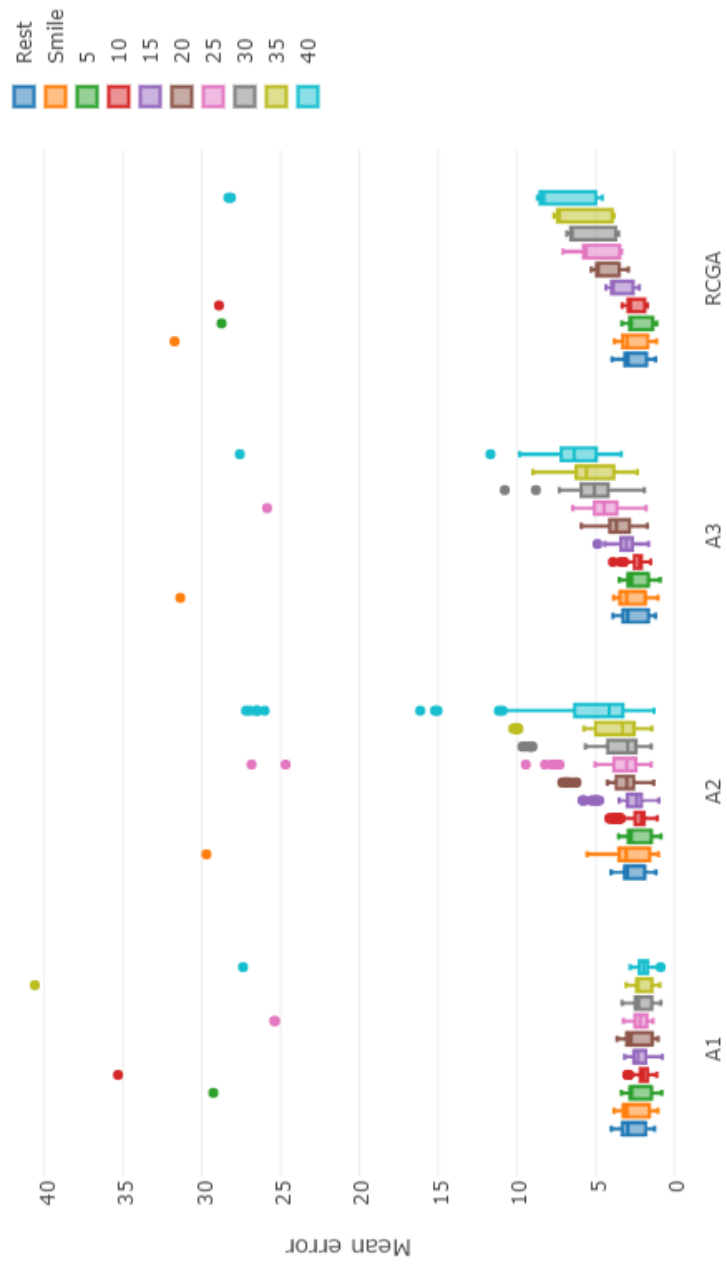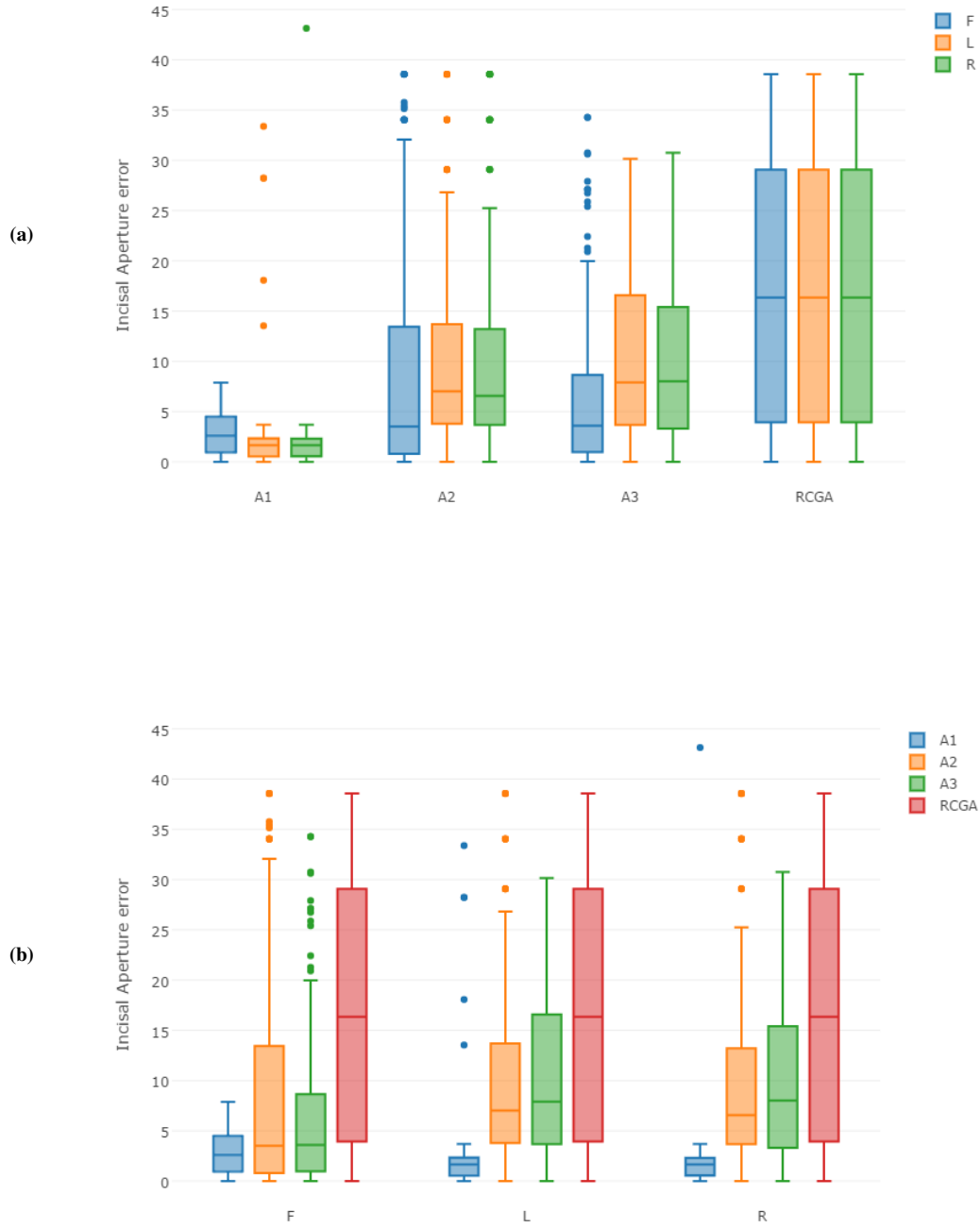Figure 10: (Continued)

(a)

(b)

Figure 11: Box-plots corresponding to inter-incisor distance error (in mm) grouped by : (a) algorithm and pose, (b) pose and algorithm, (c) algorithm and mandible aperture.
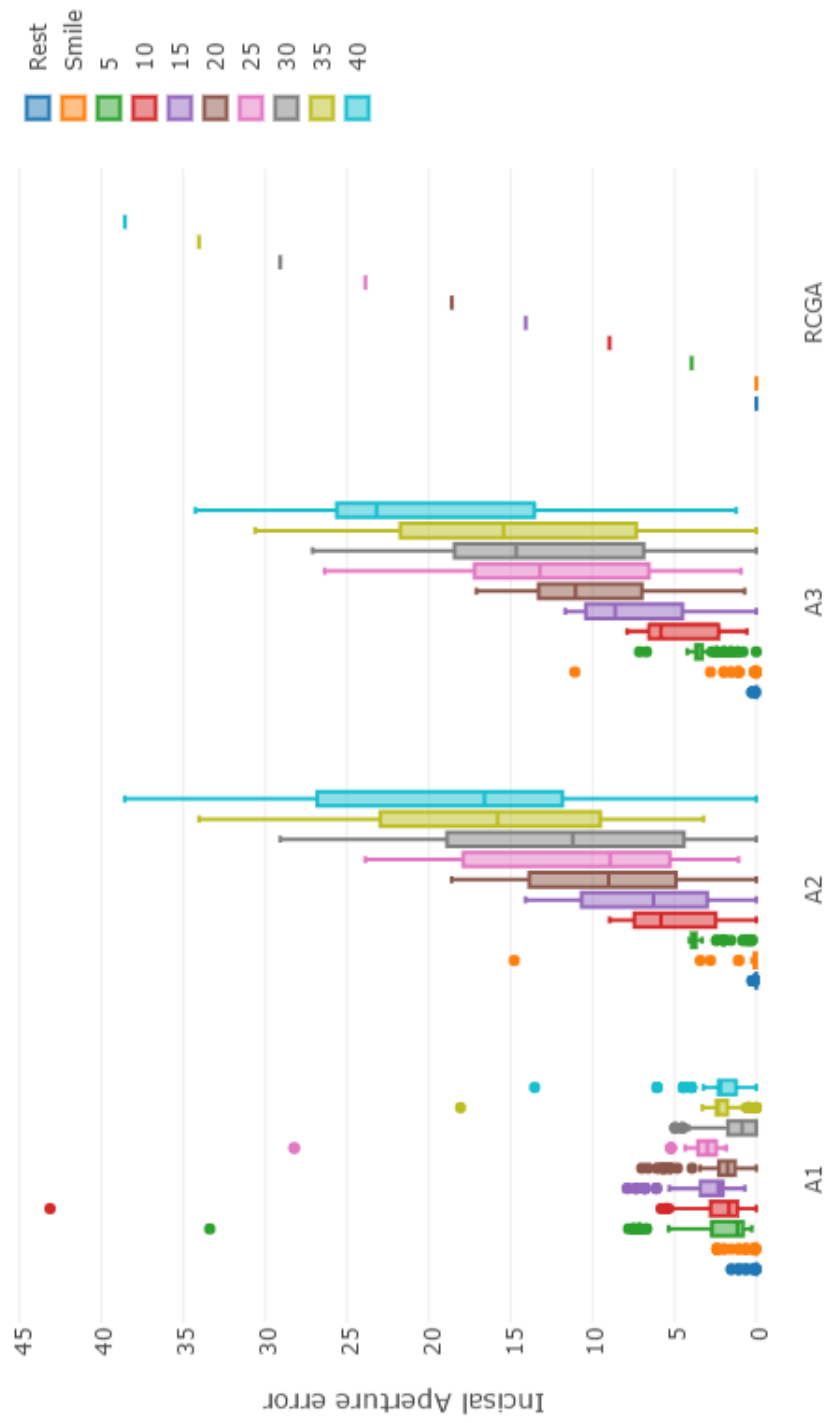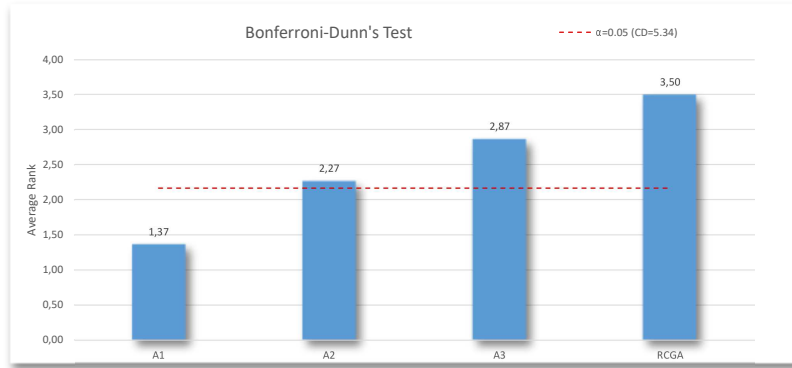
14

Figure 11: (Continued)

15

Figure 12: Bonferroni-Dunn's graphic corresponding to the results.

In summary, the visual analysis corroborates the previous numerical analysis, clearly showing the substantial improvement of A1 over RCGA in terms of quality of the results. A1 obtains accurate superimpositions particularly when the mandibular aperture is greater than 10 *mm* (this effect is distinctly visible in Fig. 14). Considering most of the instances, RCGA manages to obtain relatively good mean error values but at the expense of introducing a considerable distortion of the skull (mostly noticeable for bigger apertures, see the last row in Fig. 13).

4.5. Discussion

Together with obvious anatomical foundations, numeric, visual, and statistical analysis of the results justify the choice to consider the mandible articulation within the automatic SFO method. In addition, these results validate the aperture model, showing significant differences between our novel proposal and the state-of-the-art method. More importantly, our best variant A1 not only reduces notably the mean error of the original approach but it also obtains good results on the estimation of the actual mandible aperture in most of the instances.

The traditional CFS technique requires a forensic expert to position the skull in the same pose as the face in the photograph [10, 37]. In previous computer-based SFO methods, two manual approaches were followed to approximate the mandible aperture [24]: 1) Before capturing the 3D model, the mandible was normally located relative to the cranium so that the model resembled the facial expression of the photograph under study: and 2) Once the mandible and the cranium were scanned, both 3D models were manually positioned according to the relative aperture in the photograph. Both these manual approaches lead to important shortcomings and they are even more significant given the multiple comparisons typically performed in CFS to check the correspondence of one skull to different missing/deceased candidates. Additionally, each individual comparison should involve the analysis of one skull against more than one AM photograph of the same person because the reliability and accuracy of the method significantly increases when proceeding in that way [14, 50, 2]. Overall, this is a very time consuming task even using an automatic SFO method.

In our proposal, the expert is only required to pinpoint the incisors, and the mandibular and cranial condyles, a much simpler and faster task to carry out. This task is only performed once, no matter which number of AM photos the skull is compared with. Thereby, our approach avoids entirely the time consuming and error-prone mandible positioning of the skull, besides adding versatility to the procedure as it can adapt the 3D model pose to different mouth openings in the photographs.

Nevertheless, we identified some limitations in our proposal. We introduced different sources of error that may affect the accuracy of the technique:

- Our main concern is the **suitability of the considered mandible aperture model**. Fig. 15 reflects the limitations of the prediction model used in this work. The figure shows a small antero-posterior displacement between the predicted and the actual position of the mandible. This displacement is increased along with the aperture percentage. Considering small apertures, the prediction error is minimal, but it is notably higher when
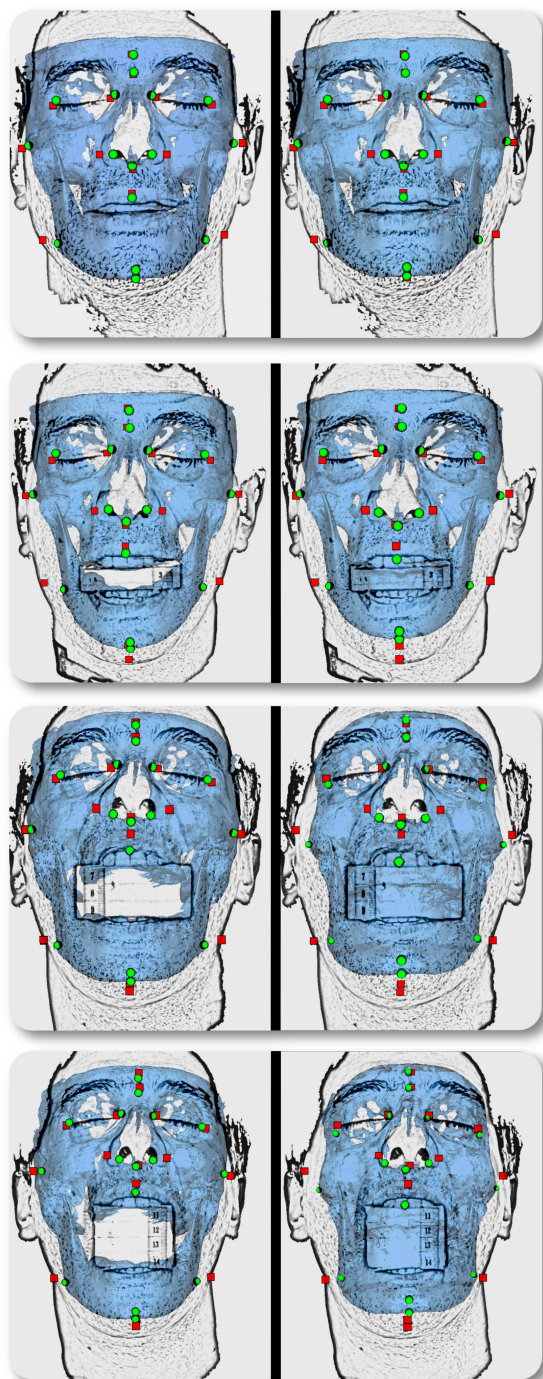
16

Figure 13: (Color online) Visual SFO results: A1 (left images) and RCGA (right images) for each of the frontal instances (Smile, 15, 30, and 40).
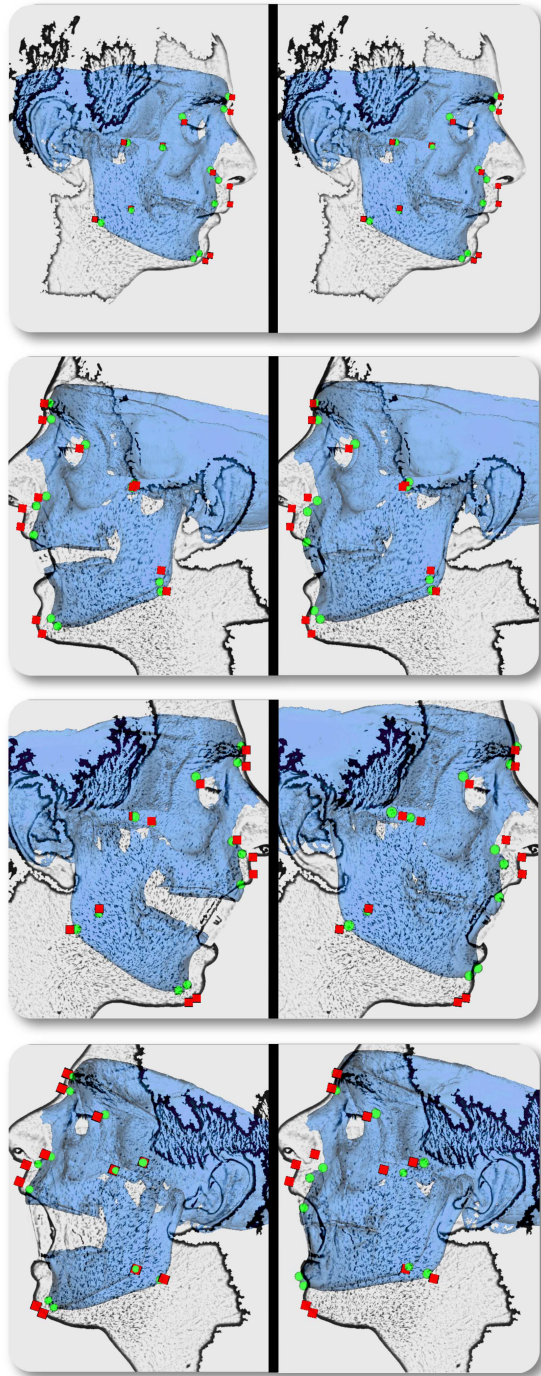
Figure 14: (Color online) Visual SFO results: A1 (left images) and RCGA (right images) for each of the lateral instances (Smile, 15, 30, and 40).

considering bigger apertures. Fortunately, most people do not usually appear in photographs with unnatural expressions (with apertures ranging > 25 *mm*) in real-world cases. As said, a practical parameterization of the problem was essential as long as the prediction model satisfied a minimum reliability standard. Therefore, the good quality of the results validates the suitability of the articulated SFO approach for a real environment, even considering a simplified aperture model. Future works can focus on developing and integrating different, more refined aperture models that can provide an even more realistic articulation of the mandible from an anatomical viewpoint.
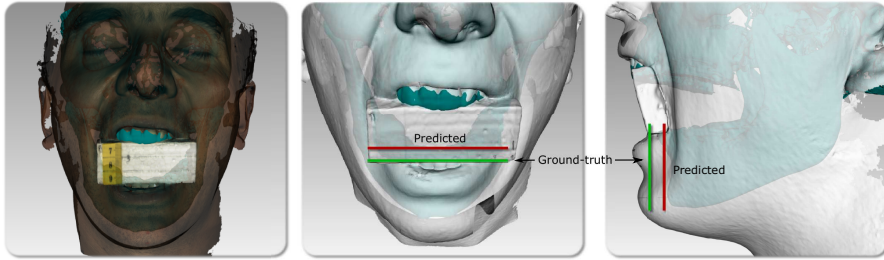

Figure 15: (Color online) Ground truth position of the aligned face and skull, including the predicted position of the mandible. Red lines indicate the erroneous predicted location of the inferior incisor by the considered model, while green lines mark its true location.

- Another source of error is the **manual alignment between 3D scans** to build the ground truth database. In previous works [5, 25], the 3D models of face and skull were perfectly aligned, as they both were segmented from CBCT scans. In our work, we introduced another source of uncertainty since for each case in the database, the skull model has to be aligned with the facial model. Moreover, each facial model comprises a different amount of soft tissue deformation related to the amount of aperture. We minimized this error by manually registering the CBCT facial model with the undistorted parts for each facial model of the database; thus, it can be considered as imperceptible.

- The SFO approach followed in this work relies on the correspondence between landmarks located in the face and the skull. The ground truth technique introduced here allows us to automatically extract the facial landmarks along with the 2D photographs from the 3D model using a known projection matrix. This technique ensures a perfect correspondence among pairs of facial points in the photograph and the 3D model, since the facial pose remains consistent. Since the facial models are aligned with the skull model, it is possible to remove potential sources of error regarding landmark placement for the SFO. However, the forensic expert is still required to **pinpoint the landmarks in the 3D models**. The task of manually placing landmarks in 3D models and photographs has already been identified as a significant source of imprecision [13, 4].

5. Concluding Remarks

Our previous works [26, 27, 5] identified different sources of uncertainty that might affect the accuracy of SFO approaches. In this contribution, we have tackled one such source in order to improve the reliability of the technique. In particular, we have integrated the mandible articulation within the SFO optimization algorithm. According to the results we obtained, the application of a simple mandible aperture model have proven its effectiveness, significantly improving the accuracy and the versatility of the state-of-the-art automatic SFO method. Such an outstanding performance facilitates the use of facial photos in which the individual either smiles or opens the mouth partially. Such photos have been usually discarded in real identification scenarios. The availability of new AM pieces of evidence is a crucial consequence of this contribution; indeed, the use of multiple facial photos of the same individual is essential to increase the reliability of the identification based in CFS [14, 50, 2].

Although the results of the aperture model presented a small error when considering large mandible apertures, it is an adequate prediction model for the SFO problem since large mandible apertures are not acceptable in real

identification cases. We must remark that such a simplified aperture model yielded in huge improvements in both the optimization process and the quality results of the SFO algorithm. Beyond the achieved improvement, one of our future works aims to analyze the use of more complex and realistic aperture models.

6. Acknowledgments

References

[1] X. Alomar, J. Medrano, J. Cabratosa, J. A. Clavero, M. Lorente, I. Serra, J. M. Monill, A. Salvador, Anatomy of the Temporomandibular Joint, Seminars in Ultrasound, CT and MRI 28 (3) (2007) 170–183. `doi:10.1053/j.sult.2007.02.002`.

[2] D. Austin-Smith, W. R. Maples, The reliability of skull/photograph superimposition in individual identification, Journal of Forensic Sciences 39 (2) (1994) 446–55.

[3] J. P. Baeyens, H. Gilomen, B. Erdmann, R. Clijsen, J. Cabri, D. Vissers, In vivo measurement of the 3D kinematics of the temporomandibular joint using miniaturized electromagnetic trackers: Technical report, Medical and Biological Engineering and Computing 51 (4) (2013) 479–484. `doi:10.1007/s11517-012-1015-4`.

[4] B. R. Campomanes-Álvarez, O. Ibáñez, F. Navarro, I. Alemán, O. Cordón, S. Damas, Dispersion assessment in the location of facial landmarks on photographs, International Journal of Legal Medicine 129 (1) (2014) 227–236. `doi:10.1007/s00414-014-1002-4`.

[5] B. R. Campomanes-Álvarez, O. Ibáñez, C. Campomanes-Álvarez, S. Damas, O. Cordón, Modeling Facial Soft Tissue Thickness for Automatic Skull-Face Overlay, IEEE Transactions on Information Forensics and Security 10 (10) (2015) 2057–2070. `doi:10.1109/TIFS.2015.2441000`.

[6] B. R. Campomanes-Álvarez, O. Ibáñez, F. Navarro, I. Alemán, M. Botella, S. Damas, O. Cordón, Computer vision and soft computing for automatic skull-face overlay in craniofacial superimposition, Forensic Science International 245 (2014) 77–86. `doi:10.1016/j.forsciint.2014.10.009`.

[7] C. Campomanes-Álvarez, O. Ibáñez, O. Cordón, Design of criteria to assess craniofacial correspondence in forensic identification based on computer vision and fuzzy integrals, Applied Soft Computing 46 (2016) 596–612. `doi:10.1016/j.asoc.2015.11.006`.

[8] C. Campomanes-Álvarez, O. Ibáñez, O. Cordón, C. Wilkinson, Hierarchical information fusion for decision making in craniofacial superimposition, Information Fusion 39 (2018) 25–40. `doi:10.1016/j.inffus.2017.03.004`.

[9] C. Cattaneo, Forensic anthropology: developments of a classical discipline in the new millennium, Forensic Science International 165 (2-3) (2007) 185–193. `doi:10.1016/j.forsciint.2006.05.018`.

[10] P. Chandra Sekharan, Positioning the Skull for Superimposition, in: Forensic Analysis of the Skull, Iscan, MY, Helmer, R. (eds), New York, 1993, pp. 105–118.

[11] J. Chen, T. R. Katona, The limitations of the instantaneous centre of rotation in joint research, Journal of Oral Rehabilitation 26 (4) (1999) 274–9. `doi:10.1046/j.1365-2842.1999.00350.x`.

[12] X. Chen, The instantaneous center of rotation during human jaw opening and its significance in interpreting the functional meaning of condylar translation, American Journal of Physical Anthropology 106 (1) (1998) 35–46. `doi:10.1002/(SICI)1096-8644(199805)106:1<35::AID-AJPA3>3.0.CO;2-C`.

[13] M. Cummaudo, M. Guerzoni, L. Marasciuolo, D. Gibelli, A. Cigada, Z. Obertovà, M. Ratnayake, P. Poppa, P. Gabriel, S. Ritz-Timme, C. Cattaneo, Pitfalls at the root of facial assessment on photographs: A quantitative study of accuracy in positioning facial landmarks, International Journal of Legal Medicine 127 (3) (2013) 699–706. `doi:10.1007/s00414-013-0850-7`.

[14] S. Damas, C. Wilkinson, T. Kahana, E. Veselovskaya, A. Abramov, R. Jankauskas, P. T. Jayaprakash, E. Ruiz, F. Navarro, M. I. Huete, E. Cunha, F. Cavalli, J. Clement, P. Lestón, F. Molinero, T. Briers, F. Viegas, K. Imaizumi, D. Humpire, O. Ibáñez, Study on the performance of different craniofacial superimposition approaches (II): Best practices proposal, Vol. 257, Forensic Sci. Int, 2015. `doi:10.1016/j.forsciint.2015.07.045`.

[15] S. Damas, O. Cordón, O. Ibáñez, J. Santamaría, I. Alemán, M. Botella, F. Navarro, Forensic identification by computer-aided craniofacial superimposition, ACM Computing Surveys 43 (4) (2011) 1–27. `doi:10.1145/1978802.1978806`.

[16] O. J. Dunn, Multiple Comparisons Among Means, Journal of the American Statistical Association 56 (293) (1961) 52–64. `doi:10.2307/2282330`.

[17] O. Faugeras, T.-D. C. Vision, A Geometric Viewpoint, Cambridge: The MIT Press, 1993.

[18] T. W. Fenton, A. N. Heard, N. J. Sauer, Skull-photo superimposition and border deaths: Identification through exclusion and the failure to exclude, Journal of Forensic Sciences 53 (1) (2008) 34–40. `doi:10.1111/j.1556-4029.2007.00624.x`.

[19] M. Friedman, A Comparison of Alternative Tests of Significance for the Problem of m Rankings, The Annals of Mathematical Statistics 11 (1) (1940) 86–92. `doi:10.1214/aoms/1177731944`.

[20] J. A. Gal, L. M. Gallo, G. Murray, I. J. Klineberg, C. W. Johnson, S. Palla, Screw Axes and Wrenches in the Study of Human Jaw, in: A Symposium Commemorating the Legacy, Works, and Life of Sir Robert Stawell Ball Upon the 100th Anniversary of A Treatise on the Theory of Screws, 2000, pp. 1–13.

[21] L. M. Gallo, G. B. Airoldi, R. L. Airoldi, S. Palla, Description of mandibular finite helical axis pathways in asymptomatic subjects, Journal of Dental Research 76 (2) (1997) 704–713. `doi:10.1177/00220345970760021201`.

[22] S. García, F. Herrera, H. U. Es, An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons, Journal of Machine Learning Research 9 (2008) 2677–2694.

[23] T. K. Goto, G. E. J. Langenbach, T. W. P. Korioth, M. Hagiwara, M. L. Tonndorf, A. G. Hannam, Functional movements of putative jaw muscle insertions, Anatomical Record 242 (2) (1995) 278–288. `doi:10.1002/ar.1092420218`.

[24] M. I. Huete, O. Ibáñez, C. Wilkinson, T. Kahana, Past, present, and future of craniofacial superimposition: Literature and international surveys, Legal Medicine 17 (4) (2015) 267–278. `doi:10.1016/j.legalmed.2015.02.001`.

[25] O. Ibáñez, F. Cavalli, B. R. Campomanes-Álvarez, C. Campomanes-Álvarez, A. Valsecchi, M. I. Huete, Ground truth data generation for skullface overlay, International Journal of Legal Medicine 129 (3) (2015) 569–581. `doi:10.1007/s00414-014-1074-1`.

[26] O. Ibáñez, L. Ballerini, O. Cordón, S. Damas, J. Santamaría, An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification, Information Sciences 179 (23) (2009) 3998–4028. `doi:10.1016/j.ins.2008.12.029`.

[27] O. Ibáñez, O. Cordón, S. Damas, J. Santamaría, Modeling the skull-face overlay uncertainty using fuzzy sets, IEEE Transactions on Fuzzy Systems 19 (5) (2011) 946–959. `doi:10.1109/TFUZZ.2011.2158220`.

[28] O. Ibáñez, A. Valsecchi, F. Cavalli, M. I. Huete, B. R. Campomanes-Álvarez, C. Campomanes-Álvarez, R. Vicente, D. Navega, A. Ross, C. Wilkinson, R. Jankauskas, K. Imaizumi, R. Hardiman, P. T. Jayaprakash, E. Ruiz, F. Molinero, P. Lestón, E. Veselovskaya, A. Abramov, M. Steyn, J. Cardoso, D. Humpire, L. Lusnig, D. Gibelli, D. Mazzarelli, D. Gaudio, F. Collini, S. Damas, Study on the criteria for assessing skull-face correspondence in craniofacial superimposition, Legal Medicine 23 (2016) 59–70. `doi:10.1016/j.legalmed.2016.09.009`.

[29] P. T. Jayaprakash, Conceptual transitions in methods of skull-photo superimposition that impact the reliability of identification: A review, Forensic Science International 246 (2015) 110–121. `doi:10.1016/j.forsciint.2014.10.043`.

[30] Q. S. Kang, D. P. Updike, E. P. Salathe, Kinematic analysis of the human temporomandibular joint, Annals of Biomedical Engineering 21 (6) (1993) 699–707. `doi:10.1007/BF02368648`.

[31] J. H. Koolstra, T. M. van Eijden, Biomechanical analysis of jaw-closing movements, Journal Of Dental Research 74 (9) (1995) 1564–1570. `doi:10.1177/00220345950740091001`.

[32] J. H. Koolstra, T. M. G. J. Van Eijden, The jaw open-close movements predicted by biomechanical modelling, Journal of Biomechanics 30 (9) (1997) 943–950. `doi:10.1016/S0021-9290(97)00058-4`.

[33] J. K. Leader, J. R. Boston, R. E. Debski, T. E. Rudy, Mandibular kinematics represented by a non-orthogonal floating axis joint coordinate system, Journal of Biomechanics 36 (2) (2003) 275–281. `doi:10.1016/S0021-9290(02)00337-8`.

[34] J. J. Lemoine, J. J. Xia, C. R. Andersen, J. Gateno, W. Buford, M. A. K. Liebschner, Geometry-Based Algorithm for the Prediction of Nonpathologic Mandibular Movement, Journal of Oral and Maxillofacial Surgery 65 (12) (2007) 2411–2417. `doi:10.1016/j.joms.2007.05.011`.

[35] D. J. Magee, Temporomandibular Joint, in: Orthopedic Physical Assessment, 5th Edition, Saunders, 2014, Ch. 4, pp. 224–251.

[36] A. Mapelli, D. Galante, N. Lovecchio, C. Sforza, V. F. Ferrario, Translation and rotation movements of the mandible during mouth opening and closing, Clinical Anatomy 22 (3) (2009) 311–318. `doi:10.1002/ca.20756`.

[37] B. A. Nickerson, P. A. Fitzhorn, S. K. Koch, M. Charney, A Methodology for Near-Optimal Computational Superimposition of Two-Dimensional Digital Facial Photographs and Three-Dimensional Cranial Surface Meshes, Journal of Forensic Sciences 36 (2) (1991) 480–500. `doi:10.1520/JFS13050J`.

[38] D. J. Ostry, E. Vatikiotis-Bateson, P. L. Gribble, An examination of the degrees of freedom of human jaw motion in speech and mastication, Journal of Speech, Language, and Hearing Research 40 (6) (1997) 1341–1351.

[39] J. Rayne, Functional anatomy of the temporomandibular joint, British Journal of Oral and Maxillofacial Surgery 25 (2) (1987) 92–99. `doi:10.1016/0266-4356(87)90002-7`.

[40] J. Santamaría, O. Cordón, S. Damas, O. Ibáñez, Tackling the coplanarity problem in 3D camera calibration by means of fuzzy landmarks: A performance study in forensic craniofacial superimposition, in: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009, pp. 1686–1693. `doi:10.1109/ICCVW.2009.5457486`.

[41] A. P. Spence, E. B. Mason, Human anatomy and physiology, The Benjamin/Cummings Pub. Co, 1987.

[42] C. N. Stephan, Craniofacial identification: techniques of facial approximation and craniofacial superimposition, in: Handbook of Forensic Anthropology and Archaeology 25 (2009) 304–321.

[43] D. H. Ubelaker, A history of smithsonian-FBI collaboration in forensic anthropology especially in regard to facial imagery, Forensic Sciience Communications 2 (4).

[44] M. B. Villamil, L. P. Nedel, C. M. D. S. Freitas, B. Macq, Simulation of the human TMJ behavior based on interdependent joints topology, Computer Methods and Programs in Biomedicine 105 (3) (2012) 217–232. `doi:10.1016/j.cmpb.2011.09.010`.

[45] C. Wilkinson, C. Rynn, Craniofacial Identification, Cambridge University Press, 2012.

[46] V. P. Willard, Tissue Engineering of the Temporomandibular Joint, Elsevier Ltd., 2011.

[47] W. L. Xu, J. E. Bronlund, J. Potgieter, K. D. Foster, O. Röhrle, A. J. Pullan, J. A. Kieser, Review of the human masticatory system and masticatory robotics, Mechanism and Machine Theory 43 (11) (2008) 1353–1375. `doi:10.1016/j.mechmachtheory.2008.06.003`.

[48] M. Yatabe, A. Zwijnenburg, C. C. Megens, M. Naeije, The kinematic center: a reference for condylar movements, Journal of Dental Research 74 (10) (1995) 1644–1648. `doi:10.1177/00220345950740100401`.

[49] M. Yoshino, Craniofacial superimposition, Craniofacial Identification, University Press, Cambridge, 2012.

[50] M. Yoshino, K. Imaizumi, S. Miyasaka, S. Seta, Evaluation of anatomical consistency in craniofacial superimposition images, Forensic Science International 74 (1-2) (1995) 125–134. `doi:10.1016/0379-0738(95)01742-2`.

Bibliography

[ACL04]   M. Andreetto, G. M. Cortelazzo, and L. Lucchese. Frequency domain registration of computer tomography data. *Proceedings - 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT)*, pp. 550–557, 2004.

[AE94]   N. Asada and M. Esashi. Planar type mirror galvanometer and method of manufacture. *Patent US5606447*, 1994.

[AFP00]   M. A. Audette, F. P. Ferrie, and T. M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201–217, 2000.

[AHB87]   K. S. Arun, T. S. Huang, and S. D. Blostein. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.

[AISB95]   W. A. Aulsebrook, M. Y. Işcan, J. H. Slabbert, and P. Becker. Superimposition and reconstruction in forensic facial identification: a survey. *Forensic Science International*, 75(2-3):101–120, 1995.

[AMC⁺07]   X. Alomar, J. Medrano, J. Cabratosa, J. A. Clavero, M. Lorente, I. Serra, J. M. Monill, and A. Salvador. Anatomy of the Temporomandibular Joint. *Seminars in Ultrasound, CT and MRI*, 28(3):170–183, 2007.

[ASM94]   D. Austin-Smith and W. R. Maples. The reliability of skull/photograph superimposition in individual identification. *Journal of Forensic Sciences*, 39(2):446–55, 1994.

[ATS⁺03]   P. Allen, A. Troccoli, B. Smith, S. Murray, I. Stamos, and M. Leordeanu. New methods for digital modeling of historic sites. *IEEE Computer Graphics and Applications*, 23(6):32–41, 2003.

[BCAV⁺17]   E. Bermejo, C. Campomanes-Álvarez, A. Valsecchi, O. Ibáñez, S. Damas, and O. Cordón. Genetic algorithms for skull-face overlay including mandible articulation. *Information Sciences*, 420:200–217, 2017.

[BCD⁺07]   L. Ballerini, O. Cordón, S. Damas, J. Santamaría, I. Aleman, and M. Botella. Craniofacial Superimposition in Forensic Identification using Genetic Algorithms. In *IAS '07: Proceedings of the Third International Symposium on Information Assurance and Security*, pp. 429–434, 2007.

[BCDS15]   E. Bermejo, O. Cordón, S. Damas, and J. Santamaría. A comparative study on the application of advanced bacterial foraging models to image registration. *Information Sciences*, 295:160–181, 2015.

[BCM05]   A. Buades, B. Coll, and J. M. Morel. A Review of Image Denoising Algorithms, with a New One. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.

[BD01]   H.-G. Beyer and K. Deb. On self-adaptive features in real-parameter evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 5(3):250–270, 2001.

[BDGG09]   L. Bianchi, M. Dorigo, L. M. Gambardella, and W. J. Gutjahr. A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing*, 8(2):239–287, 2009.

[BDW86]   R. J. Bastiaan, G. D. Dalitz, and C. Woodward. Video superimposition of skulls and photographic portraits-a new aid to identification. *Journal of Forensic Science*, 31:1373–1379, 1986.

[Ber96]   A. Bertillon. *The Bertillon system of identification*. Chicago, 1896.

[BETV08]   H. Bay, a. Ess, T. Tuytelaars, and L. Vangool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359, 2008.

[BFM97]   T. Bäck, D. B. Fogel, and Z. Michalewicz. *Handbook of Evolutionary Computation*, vol. 2. IOP Publishing Ltd. Bristol, 1997.

[BGE$^+$13]   J. P. Baeyens, H. Gilomen, B. Erdmann, R. Clijsen, J. Cabri, and D. Vissers. In vivo measurement of the 3D kinematics of the temporomandibular joint using miniaturized electromagnetic trackers: Technical report. *Medical and Biological Engineering and Computing*, 51(4):479–484, 2013.

[BH89]   L. M. Brocklebank and C. J. Holmgren. Development of equipment for the standardization of skull photographs in personal identifications by photographic superimposition. *Journal of Forensic Science*, 34(5):1214–1221, 1989.

[BHB01]   C. Baillard, P. Hellier, and C. Barillot. Segmentation of brain 3D MR images using level sets and dense registration. *Medical Image Analysis*, 5(3):185–194, 2001.

[BHMB10]   M. Böhme, M. Haker, T. Martinetz, and E. Barth. Shading constraint improves accuracy of time-of-flight measurements. *Computer Vision and Image Understanding*, 114(12):1329–1335, 2010.

[Bis89]   J. M. Bishop. Stochastic searching networks. *Artificial Neural Networks, First IEE International Conference on*, pp. 329–331, 1989.

[BL95]   G. Blais and M. Levine. Registering Multiview Range Data to Create 3D Computer Objects Registering Multiview Range Data to Create 3D Computer Objects. *Evaluation*, 17(514):820–824, 1995.

[BLR03]   L. Bornaz, A. Lingua, and F. Rinaudo. Multiple Scan Registration In Lidar Close-range Applications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Part 5/W12*, 14:72–77, 2003.

[BM92]   P. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[BR02]   F. Bernardini and H. Rushmeier. The 3D model acquisition pipeline. In *Computer Graphics Forum*, vol. 21, pp. 149–172. Blackwell Publishers Ltd, 2002.

[BR03] C. Blum and A. Roli. Metaheuristics in combinatorial optimization. *ACM Computing Surveys*, 35(3):268–308, 2003.

[Bro75] P. Broca. *Instructions craniologiques et craniométriques de la Société d'Anthropologie de Paris*, vol. 16. 1875.

[Bro92] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[BS96] K. Brunnstrom and A. Stoddart. Genetic algorithms for free-form surface matching. In *Proceedings of 13th International Conference on Pattern Recognition*, pp. 689–693 vol.4. IEEE, 1996.

[BSA05] H. Biwasaka, K. Saigusa, and Y. Aoki. The applicability of holography in forensic identification: a fusion of the traditional optical technique and digital technique. *Journal of forensic sciences*, 50(2):393–399, 2005.

[BSH+13] S. Bauer, A. Seitel, H. G. Hofmann, T. Blum, J. Wasza, M. Balda, H.-P. Meinzer, N. Navab, J. Hornegger, and L. Maier-Hein. Real-Time Range Imaging in Health Care: A Survey. In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, pp. 228–254. Springer, Berlin, Heidelberg, 2013.

[BT01] T. Butz and J. P. Thiran. Affine registration with feature space mutual information. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2208:549–556, 2001.

[BTBW77] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching: two new techniques for image matching. In *Proceedings of the 5th international joint conference on Artificial intelligence - Volume 2*, pp. 659–663. Morgan Kaufmann Publishers Inc., 1977.

[Bur99] K. Burns. *Forensic anthropology training manual*. Prentice Hall, 1999.

[CÁCÁG+17] C. Campomanes-Álvarez, B. R. Campomanes-Álvarez, S. Guadarrama, O. Ibáñez, and O. Cordón. An experimental study on fuzzy distances for skull–face overlay in craniofacial superimposition. *Fuzzy Sets and Systems*, 318:100–119, 2017.

[CÁIC16] C. Campomanes-Álvarez, O. Ibáñez, and O. Cordón. Design of criteria to assess craniofacial correspondence in forensic identification based on computer vision and fuzzy integrals. *Applied Soft Computing*, 46:596–612, 2016.

[CÁICW18] C. Campomanes-Álvarez, O. Ibáñez, O. Cordón, and C. Wilkinson. Hierarchical information fusion for decision making in craniofacial superimposition. *Information Fusion*, 39:25–40, 2018.

[CÁIN+14a] B. R. Campomanes-Álvarez, O. Ibáñez, F. Navarro, I. Alemán, O. Cordón, and S. Damas. Dispersion assessment in the location of facial landmarks on photographs. *International Journal of Legal Medicine*, 129(1):227–236, 2014.

[CÁIN+14b] B. R. Campomanes-Álvarez, O. Ibáñez, F. Navarro, I. Alemán, M. Botella, S. Damas, and O. Cordón. Computer vision and soft computing for automatic skull-face overlay in craniofacial superimposition. *Forensic Science International*, 245:77–86, 2014.

[CÁInCA+15] B. Campomanes-Álvarez, O. Ibáñez, C. Campomanes-Álvarez, S. Damas, and
O. Cordón. Modeling Facial Soft Tissue Thickness for Automatic Skull-Face Overlay.
IEEE Transactions on Information Forensics and Security, 10(10):2057–2070, 2015.

[CCFM08] U. Castellani, M. Cristani, S. Fantoni, and V. Murino. Sparse points matching by
combining 3D mesh saliency with statistical descriptors. *Computer Graphics Forum*,
27(2):643–652, 2008.

[CDB03] O. Cordón, S. Damas, and E. Bardinet. 2D Image registration with iterated local
search. In *Advances in Soft Computing*, pp. 233–242. Springer London, London,
2003.

[CDS03] O. Cordón, S. Damas, and J. Santamaría. A CHC evolutionary algorithm for 3D
image registration. *Fuzzy Sets and Systems - Ifsa 2003, Proceedings*, 2715:404–411,
2003.

[CDS06a] O. Cordón, S. Damas, and J. Santamaría. A fast and accurate approach for 3D image
registration using the scatter search evolutionary algorithm. *Pattern Recognition
Letters*, 27(11):1191–1200, 2006.

[CDS06b] O. Cordón, S. Damas, and J. Santamaría. Feature-based image registration by means
of the CHC evolutionary algorithm. *Image and Vision Computing*, 24(5):525–533,
2006.

[CDSM08] O. Cordón, S. Damas, J. Santamaría, and R. Martí. Scatter Search for the Point-
Matching Problem in 3D Image Registration. *INFORMS Journal on Computing*,
20(1):55–68, 2008.

[CE97] D. L. Collins and A. C. Evans. Animal: Validation and applications of nonlinear
registration-based segmentation. *International Journal Of Pattern Recognition And
Artificial Intelligence*, 11(8):1271–1294, 1997.

[CEGL01] P. Chalermwat, T. El-Ghazawi, and J. Lemoigne. 2-Phase GA-based image
registration on parallel clusters. *Future Generation Computer Systems*, 17(4):467–
476, 2001.

[CF01] R. J. Campbell and P. J. Flynn. A Survey Of Free-Form Object Representation and
Recognition Techniques. *Computer Vision and Image Understanding*, 81(2):166–
210, 2001.

[CGM+13] M. Cummaudo, M. Guerzoni, L. Marasciuolo, D. Gibelli, A. Cigada, Z. Obertovà,
M. Ratnayake, P. Poppa, P. Gabriel, S. Ritz-Timme, and C. Cattaneo. Pitfalls at
the root of facial assessment on photographs: A quantitative study of accuracy in
positioning facial landmarks. *International Journal of Legal Medicine*, 127(3):699–
706, 2013.

[Cha93] P. Chandra Sekharan. Positioning the Skull for Superimposition. In *Forensic
Analysis of the Skull*, pp. 105–118. Iscan, MY, Helmer, R. (eds), New York, 1993.

[Che98] X. Chen. The instantaneous center of rotation during human jaw opening and its
significance in interpreting the functional meaning of condylar translation. *American
Journal of Physical Anthropology*, 106(1):35–46, 1998.

[CK99] J. Chen and T. R. Katona. The limitations of the instantaneous centre of rotation
in joint research. *Journal of Oral Rehabilitation*, 26(4):274–279, 1999.

[CKA+12]   D. D. B. Carvalho, S. Klein, Z. Akkus, G. L. Ten Kate, H. Tang, M. Selwaness, A. F. L. Schinkel, J. G. Bosch, A. Van Der Lugt, and W. J. Niessen. Registration of free-hand ultrasound and MRI of carotid arteries through combination of point-based and intensity-based algorithms. In *Biomedical Image Registration: 5th International Workshop, WBIR 2012*, vol. 7359 LNCS, pp. 131–140. Springer, Berlin, Heidelberg, 2012.

[CM92]   Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

[CMD+95]   A. Collignon, F. Maes, D. Delaere, D. Van der Meulen, P. Suetens, and G. Marchal. Automated Multi-Modality Image Registration based on Information Theory. *Information Processing in Medical Imaging*, 1995.

[CMG+10]   P. Coupé, J. V. Manjón, E. Gedamu, D. Arnold, M. Robles, and D. L. Collins. Robust Rician noise estimation for MR images. *Medical Image Analysis*, 14(4):483–493, 2010.

[CS05]   C. Chao and I. Stamos. Semi-automatic range to range registration: A feature-based method. In *Proceedings of International Conference on 3-D Digital Imaging and Modeling, 3DIM*, pp. 254–261. IEEE, 2005.

[CS10]   S. Cui and Y. C. Soh. Linearity Indices and Linearity Improvement of 2-D Tetralateral Position-Sensitive Detector. *IEEE Transactions on Electron Devices*, 57(9):2310–2316, 2010.

[CTL04]   C. K. Chow, H. T. Tsui, and T. Lee. Surface registration using a dynamic genetic algorithm. *Pattern Recognition*, 37:105–117, 2004.

[CWS+03]   H. Chui, L. Win, R. Schultz, J. S. Duncan, and A. Rangarajan. A unified non-rigid feature registration method for brain mapping. *Medical Image Analysis*, 7(2):113–130, 2003.

[DBDA09]   S. Das, A. Biswas, S. Dasgupta, and A. Abraham. Bacterial foraging optimization algorithm: theoretical foundations, analysis, and applications. *Foundations of Computational Intelligence*, 3:23–55, 2009.

[DC18]   S. Damas and O. Cordón. *Handbook on Craniofacial Superimposition.* Springer-Verlag New York Inc, 2018.

[DCI+11]   S. Damas, O. Cordón, O. Ibáñez, J. Santamaría, I. Alemán, M. Botella, and F. Navarro. Forensic identification by computer-aided craniofacial superimposition. *ACM Computing Surveys*, 43(4):1–27, 2011.

[DCS11]   S. Damas, O. Cordón, and J. Santamaría. Medical Image Registration Using Evolutionary Computation: An Experimental Survey. *IEEE Computational Intelligence Magazine*, 6(4):26–42, 2011.

[DCV+86]   P. Delfino, M. Colonna, E. Vacca, F. Potente, and F. Introna. Computer-aided skull/face superimposition. *The American journal of forensic medicine and pathology*, 7(3):201–12, 1986.

[DDAB09]   S. Dasgupta, S. Das, A. Abraham, and A. Biswas. Adaptive Computational Chemotaxis in Bacterial Foraging Optimization: An Analysis. *IEEE Transactions on Evolutionary Computation*, 13(4):919–941, 2009.

[DDMT08]   I. De Falco, A. Della Cioppa, D. Maisto, and E. Tarantino. Differential Evolution as a viable tool for satellite image registration. *Applied Soft Computing Journal*, 8(4):1453–1462, 2008.

[DF01]   G. Dalley and P. Flynn. Range image registration: A software platform and empirical evaluation. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pp. 246–253. IEEE Computer Society, 2001.

[DG14]   S. Devi and M. Geethanjali. Application of Modified Bacterial Foraging Optimization algorithm for optimal placement and sizing of Distributed Generation. *Expert Systems with Applications*, 41(6):2772–2781, 2014.

[DM03]   L. Di Stefano and S. Mattoccia. Fast template matching using bounded partial correlation. *Machine Vision and Applications*, 13(4):213–221, 2003.

[DS04]   M. Dorigo and T. Stützle. *Ant colony optimization*. MIT Press, 2004.

[Dun61]   O. J. Dunn. Multiple Comparisons Among Means. *Journal of the American Statistical Association*, 56(293):52–64, 1961.

[DWK+15]   S. Damas, C. Wilkinson, T. Kahana, E. Veselovskaya, A. Abramov, R. Jankauskas, P. T. Jayaprakash, E. Ruiz, F. Navarro, M. I. Huete, E. Cunha, F. Cavalli, J. Clement, P. Lestón, F. Molinero, T. Briers, F. Viegas, K. Imaizumi, D. Humpire, and O. Ibáñez. *Study on the performance of different craniofacial superimposition approaches (II): Best practices proposal*, vol. 257. Forensic Science International, 2015.

[EJ16]   A. Elangovan and T. Jeyaseelan. Medical imaging modalities: A survey. In *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*, pp. 1–4. IEEE, 2016.

[EOHM10]   T. Edeler, K. Ohliger, S. Hussmann, and A. Mertins. Time-of-flight depth image denoising using prior noise information. *Signal Processing Proceedings (ICSP)*, pp. 119–122, 2010.

[ES93]   L. J. Eshelman and J. D. Schaffer. Real-coded genetic algorithms and interval schemata. In D. L. Whitley, editor, *Foundation of Genetic Algorithms 2*, pp. 187–202, San Mateo, CA, 1993. Morgan Kaufmann.

[ES03]   A. E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing*. Natural Computing Series. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[Esh91]   L. J. Eshelman. The CHC Adaptive Search Algorithm: How to Safe Search When Engaging in Nontraditional Genetic Recombination. *Foundations of genetic algorithms*, pp. 265–283, 1991.

[FBC13]   G. M. Farinella, S. Battiato, and R. Cipolla. *Advanced Topics in Computer Vision*. Springer, 2013.

[FF98]   P. A. Freeborough and N. C. Fox. Modeling brain deformations in Alzheimer disease by fluid registration of serial 3D MR images. *Journal of Computer Assisted Tomography*, 22(5):838–843, 1998.

[FG13]   G. Fornarelli and A. Giaquinto. An unsupervised multi-swarm clustering technique for image segmentation. *Swarm and Evolutionary Computation*, 11:31–45, 2013.

[FGVg84]   J. M. Fitzpatrick, J. J. Grefenstette, and D. Van-gucht. Image registration by genetic search. *IEEE Southeast Conference*, 1984.

[FH08]   S. Fuchs and G. Hirzinger. Extrinsic and depth calibration of ToF-cameras. In *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2008.

[FHS08]   T. W. Fenton, A. N. Heard, and N. J. Sauer. Skull-photo superimposition and border deaths: Identification through exclusion and the failure to exclude. *Journal of Forensic Sciences*, 53(1):34–40, 2008.

[FM96]   L. M. G. Fonseca and B. S. Manjunath. Registration Techniques for Multisensor Remotely Sensed Imagery. *Photogrammetric Engineering & Remote Sensing*, 62(September):1049–1056, 1996.

[FM08]   S. Fuchs and S. May. Calibration and registration for precise surface reconstruction with Time-Of-Flight cameras. *International Journal of Intelligent Systems Technologies and Applications*, 5(3-4), 2008.

[Fog06]   D. B. Fogel. *Evolutionary computation : toward a new philosophy of machine intelligence*. John Wiley & Sons, 2006.

[FR89]   T. A. Feo and M. G. C. Resende. A probabilistic heuristic for a computationally difficult set covering problem. *Operations Research Letters*, 8(2):67–71, 1989.

[Fri40]   M. Friedman. A Comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1):86–92, 1940.

[FS94]   J. Flusser and T. Suk. A Moment-Based Approach to Registration of Images with Affine Geometric Distortion. *IEEE Transactions on Geoscience and Remote Sensing*, 32(2):382–387, 1994.

[FV93]   O. Faugeras and T.-D. C. Vision. *A Geometric Viewpoint*. Cambridge: The MIT Press, 1993.

[FYBF13]   I. Fister, X. S. Yang, J. Brest, and D. Fister. A brief review of nature-inspired algorithms for optimization. In *Elektrotehniski Vestnik/Electrotechnical Review*, vol. 80, pp. 116–122. 2013.

[GAAP97]   L. M. Gallo, G. B. Airoldi, R. L. Airoldi, and S. Palla. Description of mandibular finite helical axis pathways in asymptomatic subjects. *Journal of Dental Research*, 76(2):704–713, 1997.

[GB37]   J. Glaister and J. C. Brash. *Medico-legal aspects of the Ruxton case*. E. and S. Livingstone, Edinburgh, UK, 1st edition, 1937.

[GDMW04]   K. A. Ganser, H. Dickhaus, R. Metzner, and C. R. Wirtz. A deformable digital brain atlas system according to Talaicrach and Tournoux. *Medical Image Analysis*, 8(1):3–22, 2004.

[Geo93]   R. M. George. Anatomical and artistic guidelines for forensic facial reconstruction. In M. Y. Iscan and R. Helmer, editors, *Forensic Analysis of the Skull*, pp. 215–227. Wiley Liss, New York, NY, 1993.

[GGM+00]   J. A. Gal, L. M. Gallo, G. Murray, I. J. Klineberg, C. W. Johnson, and S. Palla. Screw Axes and Wrenches in the Study of Human Jaw mechanics. In *A Symposium Commemorating the Legacy, Works, and Life of Sir Robert Stawell Ball. Cambridge, UK; University of Cambridge, Trinity College*, pp. 1–13, 2000.

[GH08]   S. García and F. Herrera. An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. *Journal of Machine Learning Research*, 9:2677–2694, 2008.

[GIV10]   A. Georgopoulos, C. Ioannidis, and A. Valanis. Assesing the Performance of a Structured Light Scanner. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2010.

[GLK+95]   T. K. Goto, G. E. J. Langenbach, T. W. P. Korioth, M. Hagiwara, M. L. Tonndorf, and A. G. Hannam. Functional movements of putative jaw muscle insertions. *Anatomical Record*, 242(2):278–288, 1995.

[Glo77]   F. Glover. Hwuristics for Integer Programming Using Surrogate Constraints. *Decision Sciences*, 8(1):156–166, 1977.

[Glo89]   F. Glover. Tabu Search — Part I. *ORSA Journal of Computing*, 1(3):190–206, 1989.

[Glo03]   F. Glover. *Handbook of metaheuristics*. Kluwer Academic Publishers, Boston, 2003.

[GM13]   A. R. Goldman and P. D. Maldjian. Reducing Radiation Dose in Body CT: A Practical Approach to Optimizing CT Protocols. *American Journal of Roentgenology*, 200(4):748–754, 2013.

[Gol89]   D. E. Goldberg. Genetic Algorithms in Search Optimization and Machine Learning. In *Addison Wesley Publisching*, p. 412. Addison-Wesley Longman Publishing Co., Inc., 1989.

[Gol07]   L. W. Goldman. Principles of CT: radiation dose and image quality. *Journal of nuclear medicine technology*, 35(4):213–25; quiz 226–8, 2007.

[Gos05]   A. Goshtasby. *2-D and 3-D Image Registration*. Wiley Interscience, 2005.

[Gos12]   A. Goshtasby. Image Registration: Principles, Tools and Methods. *Advances in Computer Vision and Pattern Recognition*, pp. 1–436, 2012.

[GPM06]   T. Glatard, X. Pennec, and J. Montagnat. Performance evaluation of grid-enabled registration algorithms using bronze-standards. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 9(2):152–60, 2006.

[GPT10]   P. Gil, J. Pomares, and F. Torres. Analysis and Adaptation of Integration Time in PMD Camera for Visual Servoing. In *2010 20th International Conference on Pattern Recognition*, pp. 311–315. IEEE, 2010.

[GS01]   A. K. Ghosh and P. Sinha. An economised craniofacial identification system. *Forensic science international*, 117(1-2):109–19, 2001.

[GTKK13]   M. Grzegorzek, C. Theobalt, R. Koch, and A. Kolb. *Time-of-Flight and Depth Imaging. Sensors, Algorithms and Applications: Dagstuhl Seminar 2012 and GCPR Workshop on Imaging New Modalities*. Springer-Verlag Berlin Heidelberg, 2013.

[GW08]   R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* Addison-Wesley, Upper Saddle River, NJ, 3rd edition, 2008.

[HE11]   G. Häusler and S. Ettl. Limitations of Optical 3D Sensors. In *Optical Measurement of Surface Topography*, pp. 23–48. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[Hel86]   R. Helmer. Identifizierung der leichenuberreste das Josef Mengele. *Archives Kriminology*, 177:130–144, 1986.

[HG77]   R. Helmer and O. Grüner. Vereinfachte Schädelidentifizierung nach dem Superprojektionsverfahren mit Hilfe einer Video-Anlage. *International Journal of Legal Medicine*, 80(3):183–187, 1977.

[HHH01]   J. V. Hajnal, D. J. Hawkes, and D. L. G. Hill. *Medical image registration.* CRC Press, 2001.

[HIWK15]   M. I. Huete, O. Ibáñez, C. Wilkinson, and T. Kahana. Past, present, and future of craniofacial superimposition: Literature and international surveys. *Legal Medicine*, 17(4):267–278, 2015.

[HL09]   G. L. Heritage and A. R. Large. Principles of 3D Laser Scanning. In *Laser Scanning for the Environmental Sciences*, pp. 21–34. Wiley-Blackwell, Oxford, UK, 2009.

[HLCH12]   M. Hansard, S. Lee, O. Choi, and R. Horaud. *Time of Flight Cameras: Principles, Methods, and Applications.* SpringerBriefs in Computer Science. Springer, London, 2012.

[HN02]   R. He and P. A. Narayana. Global optimization of mutual information: Application to three-dimensional retrospective registration of magnetic resonance images. *Computerized Medical Imaging and Graphics*, 26(4):277–292, 2002.

[Hol92]   J. H. Holland. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, vol. 1. MIT Press Cambridge, 1992.

[Hor87]   B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629, 1987.

[HPO+05]   S. Heger, F. Portheine, J. A. K. Ohnsorge, E. Schkommodau, and K. Radermacher. User-interactive registration of bone with A-mode ultrasound. *IEEE Engineering in Medicine and Biology Magazine*, 24(2):85–95, 2005.

[HRG+09]   X. Huang, J. Ren, G. Guiraudon, D. Boughner, and T. M. Peters. Rapid dynamic image registration of the beating heart for diagnosis and surgical navigation. *IEEE Transactions on Medical Imaging*, 28(11):1802–1814, 2009.

[HS52]   M. R. Hestenes and E. Stiefel. Methods of Conjugate Gradients for Solving Linear Systems I. *Journal of Research of the National Bureau of Standards*, 49(6), 1952.

[HS88]   C. Harris and M. Stephens. A Combined Corner and Edge Detector. In *Procedings of the Alvey Vision Conference 1988*, pp. 147–151, 1988.

[Hub81]   P. J. Huber. *Robust Statistics.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, 1981.

[Hyl92]   W. L. Hylander. Functional Anatomy of the Temporomandibular Joint. *The Temporomandibular Joint: A Biological Basis for Clinical Practice*, 25(2):60–92, 1992.

[IBC⁺09]   O. Ibáñez, L. Ballerini, O. Cordón, S. Damas, and J. Santamaría. An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification. *Information Sciences*, 179(23):3998–4028, 2009.

[ICCÁ⁺15]   O. Ibáñez, F. Cavalli, B. R. Campomanes-Álvarez, C. Campomanes-Álvarez, A. Valsecchi, and M. I. Huete. Ground truth data generation for skull–face overlay. *International Journal of Legal Medicine*, 129(3):569–581, 2015.

[ICD12]   O. Ibáñez, O. Cordón, and S. Damas. A cooperative coevolutionary approach dealing with the skull-face overlay uncertainty in forensic identification by craniofacial superimposition. *Soft Computing*, 16(5):797–808, 2012.

[ICDS11]   O. Ibáñez, O. Cordón, S. Damas, and J. Santamaría. Modeling the skull-face overlay uncertainty using fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 19(5):946–959, 2011.

[Ind09]   E. Indriati. Historical perspectives on forensic anthropology in Indonesia. In S. Blau and D. H. Ubelaker, editors, *Handbook of Forensic Anthropology and Archaeology*, pp. 115–125. Left Coast Press, Walnut Creek, CA, 2009.

[Isc81]   M. Iscan. Integral forensic anthropology. *Practicing Anthropology*, 3(4):13–30, 1981.

[IVC⁺16]   O. Ibáñez, A. Valsecchi, F. Cavalli, M. I. Huete, B. R. Campomanes-Álvarez, C. Campomanes-Álvarez, R. Vicente, D. Navega, A. Ross, C. Wilkinson, R. Jankauskas, K. Imaizumi, R. Hardiman, P. T. Jayaprakash, E. Ruiz, F. Molinero, P. Lestón, E. Veselovskaya, A. Abramov, M. Steyn, J. Cardoso, D. Humpire, L. Lusnig, D. Gibelli, D. Mazzarelli, D. Gaudio, F. Collini, and S. Damas. Study on the criteria for assessing skull-face correspondence in craniofacial superimposition. *Legal Medicine*, 23:59–70, 2016.

[Jay15]   P. T. Jayaprakash. Conceptual transitions in methods of skull-photo superimposition that impact the reliability of identification: A review. *Forensic Science International*, 246:110–121, 2015.

[JH00]   B. Jähne and H. Haußecker. *Computer Vision and Applications*. Academic Press, 2000.

[JJCJ16]   S. Y. Ji, B. K. Jeong, S. Choi, and D. H. Jeong. A multi-level intrusion detection method for abnormal network behaviors. *Journal of Network and Computer Applications*, 62(TR06):9–17, 2016.

[JPS93]   D. R. Jones, C. D. Perttunen, and B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181, 1993.

[JR65]   J.A. Nelder and R.A. Mead. A Simplex method for Function Minimization. *The Computer Journal*, (7):313–380, 1965.

[JS01]   M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143–156, 2001.

[JSA01]   P. T. Jayaprakash, G. J. Srinivasan, and M. G. Amravaneswaran. Cranio-facial morphanalysis: a new method for enhancing reliability while identifying skulls by photo superimposition. *Forensic Science International*, 117(1):121–143, 2001.

[KA15]   D. P. Kumar and S. Archana. Ground Water Prediction Using Bacterial Foraging Optimization Technique. *Journal of Computer Sciences and Applications*, 3(6):172–176, 2015.

[KDH$^+$17]   M. D. Ketcha, T. De Silva, R. Han, A. Uneri, J. Goerres, M. Jacobson, S. Vogt, G. Kleinszig, and J. H. Siewerdsen. Effects of Image Quality on the Fundamental Limits of Image Registration Accuracy. 36(10):1997–2009, 2017.

[KE95]   J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of the 1995 IEEE International Conference on Neural Networks*, pp. 4:1942–1948, 1995.

[KEP99]   R. K. S. Kwan, A. C. Evans, and B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097, 1999.

[KES01]   J. F. Kennedy, R. C. Eberhart, and Y. Shi. *Swarm intelligence*. Morgan Kaufmann Publishers, 2001.

[KGV83]   S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

[Koe82]   T. D. Koelmeyer. Videocamera superimposition and facial reconstruction as an aid to identification. *The American Journal of Forensic Medicine and Pathology*, 3(1):45–48, 1982.

[Koz92]   J. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection (Complex Adaptive Systems)*. MIT Press, 1992.

[KPSV09] S. Klein, J. P. Pluim, M. Staring, and M. A. Viergever. Adaptive stochastic gradient descent optimisation for image registration. *International Journal of Computer Vision*, 81(3):227–239, 2009.

[Kri14]   S. Krig. Interest Point Detector and Feature Descriptor Survey. *Computer Vision Metrics*, (1):217–282, 2014.

[KRM$^+$10]   A. P. King, K. S. Rhode, Y. Ma, C. Yao, C. Jansen, R. Razavi, and G. P. Penney. Registering preprocedure volumetric images with intraprocedure 3-D ultrasound using an ultrasound imaging model. *IEEE Transactions on Medical Imaging*, 29(3):924–937, 2010.

[KSP07]   S. Klein, M. Staring, and J. P. W. Pluim. Evaluation of Optimization Methods for Nonrigid Medical Image Registration Using Mutual Information and B-Splines. *IEEE Transactions on Image Processing*, 16(12):2879–2890, 2007.

[KUS93]   Q. Kang, D. Updike, and E. Salathe. Kinematic analysis of the human temporomandibular joint. *Annals of Biomedical Engineering*, 21(6):699–707, 1993.

[KV97]   J. H. Koolstra and T. M. G. J. Van Eijden. The jaw open-close movements predicted by biomechanical modelling. *Journal of Biomechanics*, 30(9):943–950, 1997.

[KvE95]   J. H. Koolstra and T. M. van Eijden. Biomechanical analysis of jaw-closing
             movements. *Journal Of Dental Research*, 74(9):1564–1570, 1995.

[LBDR03]   J. K. Leader, J. R. Boston, R. E. Debski, and T. E. Rudy. Mandibular kinematics
             represented by a non-orthogonal floating axis joint coordinate system. *Journal of
             Biomechanics*, 36(2):275–281, 2003.

[LC85]   Y. Lan and D. Cai. Study on Model {TLGA-1} Skull Identification Apparatus. In
             Special Issue on Criminal Technology Supplement, p. 23. The Fifth Bureau of the
             National Public Security Department, Beijing, China, 1985.

[LCE06]   E. Lomonosov, D. Chetverikov, and A. Ekárt. Pre-registration of arbitrarily oriented
             3D surfaces using a genetic algorithm. *Pattern Recognition Letters*, 27(11):1201–
             1208, 2006.

[LCH+13]   L. Liu, P. K. Commean, C. Hildebolt, D. Sinacore, F. Prior, J. P. Carson,
             I. Kakadiaris, and T. Ju. Automated, Foot-Bone Registration Using Subdivision-
             Embedded Atlases for Spatial Mapping of Bone Mineral Density. *Journal of Digital
             Imaging*, 26(3):554–562, 2013.

[Lev44]   K. Levenberg. A method for the solution of certain non-linear problems in least
             squares. In *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168. Brown University,
             1944.

[Lew95]   J. P. Lewis. Fast Template Matching. *Canadian Image Processing and Pattern
             Recognition Society*, pp. 120–123, 1995.

[Lin13]   T. Lindeberg. Scale Selection Properties of Generalized Scale-Space Interest Point
             Detectors. *Journal of Mathematical Imaging and Vision*, 46(2):177–210, 2013.

[LKKB+15]   B. Loutfi-Krauss, J. Köhn, N. Blümer, K. Freundl, T. Koch, E. Kara, C. Scherf,
             C. Rödel, U. Ramm, and J. Licher. Effect of dose reduction on image registration and
             image quality for cone-beam CT in radiotherapy. *Strahlentherapie und Onkologie*,
             191(2):192–200, 2015.

[LKS+13]   F. Lenzen, K. I. Kim, H. Schäfer, R. Nair, S. Meister, F. Becker, C. S. Garbe,
             and C. Theobalt. Denoising strategies for time-of-flight data. In *Lecture Notes in
             Computer Science (including subseries Lecture Notes in Artificial Intelligence and
             Lecture Notes in Bioinformatics)*, vol. 8200 LNCS, pp. 25–45, 2013.

[LLH00]   J. Luck, C. Little, and W. Hoff. Registration of range data using a hybrid
             simulated annealing and iterative closest point algorithm. In *Proceedings 2000
             ICRA. Millennium Conference. IEEE International Conference on Robotics and
             Automation. Symposia Proceedings (Cat. No.00CH37065)*, vol. 4, pp. 3739–3744.
             IEEE, 2000.

[LM03]   M. Laguna and R. Marti. Scatter search: Methodology and implementations in C.
             Operations Research/ Computer Science Interfaces Series, 24:1–283, 2003.

[LMS03]   H. Lourenço, O. Martin, and T. Stützle. Iterated Local Search. *Handbook of
             Metaheuristics SE - 11*, 57:320–353, 2003.

[Low04]   D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Interna-
             tional Journal of Computer Vision*, 2004.

[LPGL16]   T. Li, Q. Pan, L. Gao, and P. Li. Differential evolution algorithm-based range image registration for free-form surface parts quality inspection. *Swarm and Evolutionary Computation*, 36:106–123, 2016.

[LRG$^+$00]   M. Levoy, S. Rusinkiewicz, M. Ginzton, J. Ginsberg, K. Pulli, D. Koller, S. Anderson, J. Shade, B. Curless, L. Pereira, J. Davis, and D. Fulk. The Digital Michelangelo Project: 3D Scanning of Large Statues. *Conference on Computer Graphics and Interactive Techniques*, pp. 131–144, 2000.

[LS10]   D. Laks and M. Sandgren. Using Time-of-Flight camera for Motion Detection and People Counting. *Master Thesis Mathematical Imaging Group Lunds University*, p. 66, 2010.

[LSD04]   T. Liu, D. Shen, and C. Davatzikos. Deformable registration of cortical structures via hybrid volumetric and surface warping. *NeuroImage*, 22(4):1790–1801, 2004.

[LSLO14]   Y. Lu, Y. Sun, R. Liao, and S. H. Ong. A pre-operative CT and non-contrast-enhanced C-arm CT registration framework for trans-catheter aortic valve implantation. *Computerized Medical Imaging and Graphics*, 38(8):683–695, 2014.

[LVJ05]   C. H. Lee, A. Varshney, and D. W. Jacobs. Mesh saliency. *ACM Transactions on Graphics*, 24(3):659, 2005.

[LXA$^+$07]   J. J. Lemoine, J. J. Xia, C. R. Andersen, J. Gateno, W. Buford, and M. A. K. Liebschner. Geometry-Based Algorithm for the Prediction of Nonpathologic Mandibular Movement. *Journal of Oral and Maxillofacial Surgery*, 65(12):2411–2417, 2007.

[LYJ03]   H. Livyatan, Z. Yaniv, and L. Joskowicz. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE Transactions on Medical Imaging*, 22(11):1395–1406, 2003.

[Maa89]   G. J. R. Maat. The positioning and magnification of faces and skulls for photographic superimposition. *Forensic Science International*, 41(3):225–235, 1989.

[Mag14]   D. J. Magee. Temporomandibular Joint. In *Orthopedic Physical Assessment*, chapter 4, pp. 224–251. Saunders, 5 edition edition, 2014.

[May09]   S. May. *3D time-of-flight ranging for robotic perception in dynamic environments*. Düsseldorf VDI-Verl, 2009.

[MBF92]   O. Monga, S. Benayoun, and O. Faugeras. From partial derivatives of 3-D density images to ridge lines. In *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1808, pp. 354–359, Champaign, Illinois, USA, 1992. IEEE.

[MBWH11]   K. Müller, S. Bauer, J. Wasza, and J. Hornegger. Automatic Multi-modal ToF / CT Organ Surface Registration. In *Bildverarbeitung für die Medizin*, pp. 154–158. Springer, Berlin, Heidelberg, 2011.

[Mey07]   J. Meyer. Histogram transformation for inter-modality image registration. *Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, BIBE*, pp. 1118–1123, 2007.

[MFP89]   V. Mandava, J. Fitzpatrick, and D. Pickens. Adaptive search space scaling in digital image registration. *IEEE Transactions on Medical Imaging*, 8(3):251–262, 1989.

[MFW⁺97]   C. R. Maurer, J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, R. J. Maciunas, and G. S. Allen. Registration of head volume images using implantable fiducial markers. *IEEE transactions on medical imaging*, 16(4):447–462, 1997.

[MGL⁺09]   A. Mapelli, D. Galante, N. Lovecchio, C. Sforza, and V. F. Ferrario. Translation and rotation movements of the mandible during mouth opening and closing. *Clinical Anatomy*, 22(3):311–318, 2009.

[MH97]   N. Mladenović and P. Hansen. Variable neighborhood search. *Computers & Operations Research*, 24(11):1097–1100, 1997.

[Mod03]   J. Modersitzki. *Numerical Methods for Image Registration*. Oxford Science Publications, 2003.

[Mos89]   P. Moscato. On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms. Technical report, 1989.

[MR13]   V. Mani and D. Rivazhagan. Survey of Medical Image Registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[MRDG02]   D. M. Muratore, J. H. Russ, B. M. Dawant, and R. L. Galloway. Three-dimensional image registration of phantom vertebrae for image-guided surgery: A preliminary study. *Computer Aided Surgery*, 7(6):342–352, 2002.

[MRH10]   A. Melbourne, G. Ridgway, and D. J. Hawkes. Image Similarity Metrics in Image Registration. In *Proceedings of SPIE - The International Society for Optical Engineering: Bellingham, US*, vol. 7623, pp. 1–10, 2010.

[MS56]   R. Martin and K. Saller. *Lehrbuch der Antropologie in systematischer Darstellung*. Gustav Fischer Verlag, Stuttgart, Germany, 1st edition, 1956.

[MV98]   J. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[MVS99]   F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.

[MW09]   S. Meers and K. Ward. Face Recognition Using a Time-of-Flight Camera. In *2009 Sixth International Conference on Computer Graphics, Imaging and Visualization*, pp. 377–382. IEEE, 2009.

[MZ04]   I. Maglogiannis and E. Zafiropoulos. Automated medical image registration using the simulated annealing algorithm. In *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, vol. 3025, pp. 456–465. Springer, Berlin, Heidelberg, 2004.

[NC09]   P. Nair and A. Cavallaro. 3-D face detection, landmark localization, and registration using a point distribution model. *IEEE Transactions on Multimedia*, 11(4):611–623, 2009.

[NFKC91]   B. A. Nickerson, P. A. Fitzhorn, S. K. Koch, and M. Charney. A Methodology for Near-Optimal Computational Superimposition of Two-Dimensional Digital Facial Photographs and Three-Dimensional Cranial Surface Meshes. *Journal of Forensic Sciences*, 36(2):480–500, 1991.

[NKDV09]   M. Nachtegael, E. Kerre, S. Damas, and D. Van der Weken. Special issue on recent advances in soft computing in image processing. *International Journal of Approximate Reasoning*, 50(1):1–2, 2009.

[NN07]   J. Novatnack and K. Nishino. Scale-Dependent 3D Geometric Features. In *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8. IEEE, 2007.

[OCL06]   G. Olague, S. Cagnoni, and E. Lutton. Introduction to the special issue on evolutionary computer vision and image understanding. *Pattern Recognition Letters*, 27(11):1161–1163, 2006.

[ODCB07]   S. Oprisescu, F. Dragos, M. Ciuc, and V. Buzuloiu. Measurements with ToF cameras and their necessary corrections. In *ISSCS 2007 - International Symposium on Signals, Circuits and Systems, Proceedings*, vol. 1, pp. 221–224. IEEE, 2007.

[OKS15]   J. J. O'Neill, T. M. Kowalewski, and R. M. Sweet. Feasibility of online semantic labeling of deformable tissues for minimally invasive surgery. *Journal of Medical Devices, Transactions of the ASME*, 9(3):9–10, 2015.

[OLZW06]   Y. S. Ong, M. H. Lim, N. Zhu, and K. W. Wong. Classification of adaptive memetic algorithms: A comparative study. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 36(1):141–152, 2006.

[OSP02]   S. Ourselin, R. Stefanescu, and X. Pennec. Robust Registration of Multi-modal Images: Towards Real-Time Clinical Applications. *Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2002*, (2489):140–147, 2002.

[OT14]   F. P. Oliveira and J. M. R. Tavares. Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2):73–93, 2014.

[OVBG97]   D. J. Ostry, E. Vatikiotis-Bateson, and P. L. Gribble. An examination of the degrees of freedom of human jaw motion in speech and mastication. *Journal of Speech, Language, and Hearing Research*, 40(6):1341–1351, 1997.

[Pas02]   K. Passino. Biomimicry of bacterial foraging for distributed optimization and control. *IEEE Control Systems Magazine*, 22(3):52–67, 2002.

[PBP08]   K. Pathak, A. Birk, and J. Poppinga. Sub-pixel depth accuracy with a time of flight sensor using multimodal Gaussian analysis. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3519–3524. IEEE, 2008.

[PLB12]   N. Pears, Y. Liu, and P. Bunting. *3D imaging, analysis and applications*. Springer, 2012.

[PMV03]   J. P. Pluim, J. B. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: A survey. 22(8):986–1004, 2003.

[Pow64]   M. J. D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, 7(2):155–162, 1964.

[Pow09]   M. J. D. Powell. The BOBYQA algorithm for bound constrained optimization without derivatives. Technical Report NA2009/06, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 2009.

[PPL+16]   D. Putzer, A. Pizzini, M. Liebensteiner, J. L. Moctezuma, and M. Nogler. Recognizing the surgical situs in minimally invasive hip arthroplasty: A comparison of different filtering techniques. *Biocybernetics and Biomedical Engineering*, 36(1):182–192, 2016.

[PT07]   G. Pascale and L. Troiano. A Niche Based Genetic Algorithm for Image Registration. *ICEIS - 9th International Conference on Enterprise Information Systems, Proceedings*, pp. 342–347, 2007.

[PXP00]   D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation 1. *Annual review of biomedical engineering*, 2(1):315–337, 2000.

[PXZF17]   Y. Pan, Y. Xia, T. Zhou, and M. Fulham. Cell image segmentation using bacterial foraging optimization. *Applied Soft Computing*, 58:770–782, 2017.

[QH07]   H. Qui and E. R. Hancock. Clustering and embedding using commute times. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1873–1890, 2007.

[RF02]   C. Robertson and R. B. Fisher. Parallel Evolutionary Registration of Range Data. *Computer Vision and Image Understanding*, 87(1-3):39–50, 2002.

[RFL02]   M. Rodrigues, R. Fisher, and Y. Liu. Special Issue on Registration and Fusion of Range Images. *Computer Vision and Image Understanding*, 87(1-3):1–7, 2002.

[RH07]   T. Ringbeck and B. Hagebeuker. A 3D time of flight camera for object detection. *Optical 3D measurement techniques*, pp. 1–10, 2007.

[RJR00]   J.-M. Rouet, J.-J. Jacq, and C. Roux. Genetic algorithms for a robust 3-D MR-CT registration. *IEEE Transactions on Information Technology in Biomedicine*, 4(2):126–136, 2000.

[RM51]   H. Robbins and S. Monro. A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 22(3):400–407, 1951.

[RMH07]   T. Ringbeck, T. Möller, and B. Hagebeuker. Multidimensional measurement by using 3-D PMD sensors. *Advances In Radio Science*, 5:135–146, 2007.

[Roh01]   K. Rohr. *Landmark-Based Image Analysis*, vol. 21 of *Computational Imaging and Vision*. Springer Netherlands, Dordrecht, 2001.

[RS13]   F. Remondino and D. Stoppa. *TOF range-imaging cameras*. Springer, 2013.

[SA15]   N. Siddique and H. Adeli. Nature Inspired Computing: An Overview and Some Future Directions. *Cognitive computation*, 7(6):706–714, 2015.

[SAB+07]   E. Stoykova, A. A. Alatan, P. Benzie, N. Grammalidis, S. Malassiotis, J. Ostermann, S. Piekh, V. Sainov, C. Theobalt, T. Thevar, and X. Zabulis. 3-D Time-Varying Scene Capture Technologies—A Survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(11):1568–1586, 2007.

[SAS13]   M. Saraswat, K. V. Arya, and H. Sharma. Leukocyte segmentation in tissue images
              using differential evolution algorithm. *Swarm and Evolutionary Computation*, 11:46–
              54, 2013.

[SBB05]   L. Silva, O. R. Bellon, and K. L. Boyer. Precision range image registration using
              a robust surface interpenetration measure and enhanced genetic algorithms. *IEEE
              Transactions on Pattern Analysis and Machine Intelligence*, 27(5):762–776, 2005.

[SBBG03]   L. Silva, O. R. P. Bellon, K. L. Boyer, and P. F. U. Gotardo. Low-Overlap Range
              Image Registration for Archaeological Applications. In *IEEE Computer Society
              Conference on Computer Vision and Pattern Recognition Workshops*, vol. 1, pp.
              9–9. IEEE, 2003.

[SCD⁺07]   J. Santamaría, O. Cordón, S. Damas, I. Aleman, and M. Botella. A scatter search-
              based technique for pair-wise 3D range image registration in forensic anthropology.
              Soft Computing, 11(9):819–828, 2007.

[SCD11]   J. Santamaría, O. Cordón, and S. Damas. A comparative study of state-of-the-art
              evolutionary image registration methods for 3D modeling. *Computer Vision and
              Image Understanding*, 115(9):1340–1354, 2011.

[SCDI09]   J. Santamaría, O. Cordón, S. Damas, and O. Ibáñez. Tackling the coplanarity
              problem in 3D camera calibration by means of fuzzy landmarks: A performance
              study in forensic craniofacial superimposition. In *2009 IEEE 12th International
              Conference on Computer Vision Workshops, ICCV Workshops*, pp. 1686–1693,
              2009.

[SDCE13]   J. Santamaría, S. Damas, O. Cordón, and A. Escámez. Self-adaptive evolution
              toward new parameter free image registration methods. *IEEE Transactions on
              Evolutionary Computation*, 17(4):545–557, 2013.

[SDGTC12]   J. Santamaría, S. Damas, J. M. García-Torres, and O. Cordón. Self-adaptive
              evolutionary image registration using differential evolution and artificial immune
              systems. *Pattern Recognition Letters*, 33(16):2065–2070, 2012.

[SDP13]   A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration:
              A survey. *IEEE Transactions on Medical Imaging*, 32(7):1153–1190, 2013.

[Sen62]   N. Sen. Identification by superimposed photographs. *International Criminal Police
              Review*, 162:284–286, 1962.

[SFPL10]   J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured
              light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, 2010.

[SG87]   G. Stockman and George. Object recognition and localization via pose clustering.
              Computer Vision, Graphics, and Image Processing, 40(3):361–387, 1987.

[Sil07]   D. Silvestre. Video surveillance using a time-of-flight camera. *Technical University
              of Denmark*, pp. 1–104, 2007.

[SKSK85]   K. Shoemake, Ken, Shoemake, and Ken. Animating rotation with quaternion curves.
              In *Proceedings of the 12th annual conference on Computer graphics and interactive
              techniques - SIGGRAPH '85*, vol. 19, pp. 245–254, New York, USA, 1985. ACM
              Press.

[SLP06]   D. Skerl, B. Likar, and F. Pernuš. A Protocol for Evaluation of Similarity Measures for Rigid Registration. *IEEE Transactions on Medical Imaging*, 25(6), 2006.

[SLP⁺07]   A. Swadzba, B. Liu, J. Penne, O. Jesorsky, and R. Kompe. A comprehensive system for 3D modeling from range images acquired from a 3D ToF sensor. In *The 5th International Conference on Computer Vision Systems (ICVS)*, 2007.

[SLW02]   G. Sharp, S. Lee, and D. Wehe. ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):90–102, 2002.

[SM87]   A. P. Spence and E. B. Mason. *Human anatomy and physiology.* The Benjamin/Cummings Pub. Co, 1987.

[SMA76]   M. Svedlow, C. McGillem, and P. Anuta. Experimental examination of similarity measures and preprocessing methods used for image registration. In P. H. Swain, D. B. Morrison, and D. E. Parks, editors, *Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Ind*, vol. 4(A), pp. 4A9–4A17, Indiana, USA, 1976.

[SMFF07]   J. Salvi, C. Matabosch, D. Fofi, and J. Forest. A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, 25(5):578–596, 2007.

[Son11]   B. C. Song. A fast normalized cross correlation-based block matching algorithm using multilevel Cauchy-Schwartz inequality. *ETRI Journal*, 33(3):401–406, 2011.

[Sör15]   K. Sörensen. Metaheuristics-the metaphor exposed. *International Transactions in Operational Research*, 22(1):3–18, 2015.

[SP97]   R. Storn and K. Price. Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[SPH01]   M. Salomon, G.-R. Perrin, and F. Heitz. Differential Evolution for Medical Image Registration. In *International Conference on Artificial Intelligence*, pp. 201–207, 2001.

[STD09]   G. Sansoni, M. Trebeschi, and F. Docchio. State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation. In *Sensors*, vol. 9, pp. 568–601. 2009.

[Sue09]   P. Suetens. *Fundamentals of Medical Imaging.* Cambridge University Press, 2009.

[SvdHK⁺09]   M. Staring, U. a. van der Heide, S. Klein, M. a. Viergever, and J. P. W. Pluim. Registration of cervical MRI using multifeature mutual information. *IEEE transactions on medical imaging*, 28(9):1412–1421, 2009.

[SW81]   F. J. Solis and R. J.-B. Wets. Minimization By Random Search Techniques. *Mathematics of Operations Research*, 6(1):19–30, 1981.

[SY93]   S. Seta and M. Yoshino. *A combined apparatus for photographic and video superimposition*, vol. 161. Wiley-Liss, New York, 1993.

[TCL⁺13]   G. K. Tam, Z. Q. Cheng, Y. K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X. F. Sun, and P. L. Rosin. Registration of 3d point clouds and meshes: A survey from rigid to Nonrigid. In *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, pp. 1199–1217, 2013.

[TM09]   B. Tso and P. M. Mather. *Classification methods for remotely sensed data.* CRC Press, 2009.

[Tsa97]   P. W. M. Tsang. A genetic algorithm for aligning object shapes. *Image and Vision Computing*, 15(11):819–831, 1997.

[TT08]   A. Tettamanzi and M. Tomassini. *Soft Computing.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[Ube99]   D. H. Ubelaker. Ales Hrdlicka's role in the history of forensic anthropology. *Journal of forensic sciences*, 44(4):724–730, 1999.

[Ube00]   D. H. Ubelaker. A history of smithsonian-FBI collaboration in forensic anthropology especially in regard to facial imagery. *Forensic Sciience Communications*, 2(4), 2000.

[UBO92]   D. H. Ubelaker, E. Bubniak, and G. O'Donnell. Computer-assisted photographic superimposition. *Journal of Forensic Science*, 37(3):750–762, 1992.

[VDLS⁺13]   A. Valsecchi, J. Dubois-Lacoste, T. Stutzle, S. Damas, J. Santamaria, and L. Marrakchi-Kacem. Evolutionary medical image registration using automatic parameter tuning. In *2013 IEEE Congress on Evolutionary Computation*, pp. 1326–1333. IEEE, 2013.

[VG10]   I. Villaverde and M. Graña. Neuro-evolutionary mobile robot egomotion estimation with a 3D ToF camera. *Neural Computing and Applications*, 20(3):345–354, 2010.

[VNFM12]   M. B. Villamil, L. P. Nedel, C. M. D. S. Freitas, and B. Macq. Simulation of the human TMJ behavior based on interdependent joints topology. *Computer Methods and Programs in Biomedicine*, 105(3):217–232, 2012.

[VW97]   P. Viola and W. I. Wells. Alignment by maximization of mutual information. In *Proceedings of IEEE International Conference on Computer Vision*, vol. 24:2, pp. 16–23, 1997.

[WEF⁺08]   X. Y. Wang, S. Eberl, M. Fulham, S. Som, and D. D. Feng. Data registration and fusion. In *Biomedical Information Technology*, pp. 187–210. Elsevier, 2008.

[Wil11]   V. P. Willard. *Tissue Engineering of the Temporomandibular Joint.* Elsevier Ltd., 2011.

[WM97]   D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997.

[WSZ⁺04]   M. Wachowiak, R. Smolikova, Y. Zheng, J. Zurada, and A. Elmaghraby. An Approach to Multimodal Biomedical Image Registration Utilizing Particle Swarm Optimization. *IEEE Transactions on Evolutionary Computation*, 8(3):289–301, 2004.

[XBP⁺08]   W. L. Xu, J. E. Bronlund, J. Potgieter, K. D. Foster, O. Röhrle, A. J. Pullan, and J. A. Kieser. Review of the human masticatory system and masticatory robotics. *Mechanism and Machine Theory*, 43(11):1353–1375, 2008.

[XD04]   Xiaoyan Xu and R. Dony. Differential evolution with powell's direction set method in medical image registration. In *2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano*, vol. 2, pp. 732–735. IEEE, 2004.

[YAF99]   S. Yamany, M. N. Ahmed, and A. A. Farag. A New Genetic-Based Technique for Matching 3-D Curves and Surfaces. *Pattern Recognition*, 32:1817–1820, 1999.

[Yan09]   X. S. Yang. Firefly algorithms for multimodal optimization. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5792, pp. 169–178. Springer, Berlin, Heidelberg, 2009.

[Yan14]   X.-S. Yang. *Nature-Inspired Optimization Algorithms*. Elsevier, 2014.

[YBS05]   S. Yoshizawa, A. Belyaev, and H. Seidel. Fast and robust detection of crest lines on meshes. *Proceedings of the 2005 ACM symposium on Solid and physical modeling*, pp. 227–232, 2005.

[YD10]   X.-S. Yang and S. Deb. Engineering Optimisation by Cuckoo Search. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(4):330–343, 2010.

[YIMS95]   M. Yoshino, K. Imaizumi, S. Miyasaka, and S. Seta. Evaluation of anatomical consistency in craniofacial superimposition images. *Forensic Science International*, 74(1-2):125–134, 1995.

[YJH13]   B. Yongqiang, X. Ji, and X. Haiyan. Practical Speech Emotion Recognition Based on Im-BFOA. *Journal of Applied Sciences*, 13(22):5349–5355, 2013.

[Yos12]   M. Yoshino. *Craniofacial superimposition*. Craniofacial Identification, University Press, Cambridge, 2012.

[YS00]   M. Yoshino and S. Seta. Skull-photo superimposition. *Encyclopedia of Forensic Sciences*, 2:807–815, 2000.

[YZMN95]   M. Yatabe, A. Zwijnenburg, C. C. Megens, and M. Naeije. The kinematic center: a reference for condylar movements. *Journal of Dental Research*, 74(10):1644–1648, 1995.

[Zad94]   L. A. Zadeh. Soft computing and fuzzy logic. *IEEE Software*, 11(6):48–56, 1994.

[Zha94]   Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.

[ZHD+12]   M. A. Zuluaga, M. Hernández Hoyos, J. C. Dávila, L. F. Uriza, and M. Orkisz. A Fast Lesion Registration to Assist Coronary Heart Disease Diagnosis in CTA Images. In *Computer Vision and Graphics*, pp. 710–717. Springer, Berlin, Heidelberg, 2012.

[Zit03]   B. Zitová. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

[ZMS+10]   O. Zvitia, A. Mayer, R. Shadmi, S. Miron, and H. K. Greenspan. Co-registration of white matter tractographies by adaptive-mean-shift and gaussian mixture modeling. *IEEE Transactions on Medical Imaging*, 29(1):132–145, 2010.

[ZSML10]   S. Zambanini, R. Sablatnig, H. Maier, and G. Langs. Automatic image-based assessment of lesion development during hemangioma follow-up examinations. *Artificial Intelligence in Medicine*, 50(2):83–94, 2010.

[ZTC99]   R. Zhang, P. Tsai, and J. Cryer. Shape-from-shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–41, 1999.

[ZW16]   W. Zhao and L. Wang. An effective bacterial foraging optimizer for global optimization. *Information Sciences*, 329:719–735, 2016.