TECHNICAL
UNIVERSITY
OF CRETE

School of Mineral Resources Engineering
Technical University of Crete,Greece

Novel and robust methods for the automatic registration of image data

Ph.D Thesis
Constantinos Spanakis

September 2020

List of Figures

List of Tables

Contents

List of Algorithms

Καινοτόμες και εύρωστες μαθηματικές τεχνικές αυτόματης ταύτισης δεδομένων εικόνων

Περίληψη

Οι διαδικασίες επεξεργασίας εικόνων συχνά περιλαμβάνουν τη σύγκριση εικόνων που απεικονίζουν τις ίδιες ή παρόμοιες σκηνές. Αυτές μπορεί να λαμβάνονται από διαφορετικές θέσεις, σε διαφορετικές χρονικές στιγμές, υπό διαφορετικές συνθήκες φωτισμού ή με τη χρήση διαφορετικών μέσων λήψης.

Για παράδειγμα, στην Ιατρική απεικόνιση, λόγω της συμπληρωματικής φύσης των απεικονιστικών μεθόδων, χρειάζεται η σύγκριση εικόνων από διαφορετικούς αισθητήρες (όπως για παράδειγμα **CT-MRI, CT- PET**) για να εκτιμηθεί η κατάσταση του ασθενούς, ώστε να αποφασιστεί το πλάνο θεραπείας. Η ταύτιση ιατρικών εικόνων με ελάχιστο σφάλμα συμβάλλει στην ακριβέστερη διάγνωση και, κατά συνέπεια, στην καλύτερη μεταχείριση και ταχύτερη ίαση του ασθενούς, αλλά και στη συντόμευση της παραμονής του στο νοσοκομείο με τη ταυτόχρονη μείωση του ιατρικού κόστους. Δορυφορικές εικόνες συχνά συγκρίνονται για την ανίχνευση φυσικών ή ανθρωπογενών αλλαγών στο περιβάλλον, επόπτευση των φυσικών πόρων ή και ανανέωση των ψηφιακών χαρτών. Σε αυτή την περίπτωση, ένα μικρό σφάλμα στην ταύτιση εικόνων τηλεπισκόπησης οδηγεί σε γεωγραφική απόκλιση πολλών χιλιομέτρων.

Παρά τη χρήση τους σε πολλούς επιστημονικούς τομείς, η Ταύτιση Εικόνων (**Image Registration**) παραμένει ένα άλυτο πρόβλημα ως προς τις μεθόδους σύγκρισης και βελτιστοποίησης. Σκοπός της ερευνητικής διαδικασίας αυτής της διδακτορικής διατριβής είναι ο σχεδιασμός καινοτόμων και εύρωστων μεθόδων για την αυτόματη ταύτιση δεδομένων εικόνων. Για αυτό το σκοπό, σχεδιάστηκαν αριθμητικές μέθοδοι προσεγγίσεων βασισμένες σε διαφορετικές μεθευρετικές μεθόδους (**metaheuristics**) σε συνδυασμό με μηχανική μάθηση και αναζήτηση ενσωμάτωσης **intensity-based** μεθόδων. Έγινε προσπάθεια να αναπτυχθεί μία ικανή αυτόματη μέθοδος που να μπορεί να χρησιμοποιηθεί για οποιοδήποτε ζεύγος εικόνων με δεδομένα από διάφορες πηγές.

Η δομή της ερευνητικής μεθοδολογίας της διατριβής περιλαμβάνει τα παρακάτω στάδια:

- Μεθευρετική Βελτιστοποίηση (**Meta-heuristic optimization**): Στη μαθηματική μοντελοποίηση του προβλήματος της ταύτισης εικόνων, θεωρείται ως ένα πρόβλημα βελτιστοποίησης, όπου η αντικειμενική συνάρτηση είναι η ομοιότητα (ανομοιότητα) των εικόνων και οι εικόνες ταυτίζονται αν αυτή η συνάρτηση μεγιστοποιείται (ελαχιστοποιείται). Γι' αυτό η μέθοδος βελτιστοποίησης πρέπει να είναι εύρωστη και ικανή να βελτιστοποιεί τη συνάρτηση με τον ελάχιστο δυνατό αριθμό επαναλήψεων.

- Υπολογιστικό κόστος: Στη Μεθευρετική Βελτιστοποίηση, οι επαναλαμβανόμενες εκτιμήσεις της αντικειμενικής συνάρτησης αυξάνουν το υπολογιστικό κόστος. Γι' αυτό, έγινε έρευνα για τη μείωση του υπολογιστικού κόστους με τη χρήση **Surrogate Models**, τα οποία κατασκευάζονται με τη χρήση μεθόδων μηχανικής μάθησης. Παρά το γεγονός ότι χρησιμοποιούνται συχνά στη Μεθευρετική Βελτιστοποίηση, τα μοντέλα αυτά δεν είχαν χρησιμοποιηθεί στην ταύτιση δεδομένων εικόνων.

- Μέτρο ομοιότητας: Το μέτρο ομοιότητας είναι ένας σημαντικός παράγοντας στην ταύτιση των εικόνων, διότι χρησιμοποιείται για τη σύγκριση τους.

Στην παρούσα έρευνα, δόθηκε έμφαση στα intensity-based measures, λόγω των πλεονεκτημάτων τους. Σκοπός είναι η εύρεση ενός τρόπου μείωσης της έκτασης των μειονεκτημάτων τους. Συγκεκριμένα, έγινε έρευνα σε μέτρα ομοιότητας βασισμένα στη στατιστική απόκλιση του Renyi (Renyi divergence), η οποία έχει χρησιμοποιηθεί ευρέως στην ταύτιση εικόνων.

Τα καινοτόμα επιτεύγματα που προκύπτουν από την ερευνητική διαδικασία της διατριβής μπορούν συνοπτικά να περιγραφούν ως:

- Βελτιστοποίηση: Στην αρχή, χρησιμοποιήθηκαν γενετικοί αλγόριθμοι, οι οποίοι έχουν χρησιμοποιηθεί ευρέως στην ταύτιση εικόνων. Συγκεκριμένα, έγινε πρόοδος στην έρευνα σχετικά με τη σχέση του ρυθμού μετάλλαξης και του αριθμού των elits στους elitist γενετικούς αλγορίθμους και πώς επηρεάζουν τη βελτιστοποίηση. Παρατηρήθηκε ότι ο αυξημένος αριθμός των elites ή/και ο αυξημένος ρυθμός μετάλλαξης συντελεί σε καλύτερη βελτιστοποίηση. Παρόλαυτα, εξαιτίας των μειονεκτημάτων τους, έγινε προσπάθεια εστίασης σε αλγορίθμους οι οποίοι δεν είχαν χρησιμοποιηθεί στην ταύτιση εικόνων μέχρι σήμερα. Η Αρμονική Αναζήτηση (Harmony Search) ήταν μία καλή εναλλακτική λόγω της απλής υλοποίησης και της ικανότητας να εκμεταλλεύεται αποτελεσματικότερα τις υποψήφιες λύσεις. Η διαδικασία αυτή δεν είχε χρησιμοποιηθεί στην ταύτιση εικόνων. Στη διάρκεια της ερευντικής διαδικασίας, αρχικά τροποποιήθηκε η μεθόδος σε μία νέα καλύτερη και ταχύτερη μέθοδο η οποία συνδυάζει αρμονική αναζήτηση με τον αλγόριθμο ALOPEX. Αυτή τη μέθοδο συγκρίναμε με μεθόδους ταύτισης του λογισμικού ITK (Insight Toolkit for Segmentation and Registration) και προέκυψε ταύτιση μεγαλύτερης ακρίβειας. Επίσης, κατασκευάσαμε ένα νέο κριτήριο τερματισμού βασισμένο στη μέση τιμή και την τυπική απόκλιση της συνάρτησης ομοιότητας των υποψηφίων λύσεων. Τέλος, επεκτείναμε τον αλγόριθμο αυτό και σχεδιάσαμε έναν νέο, ταχύτερο αλγόριθμο, ο οποίος εκμεταλλεύεται αποτελεσματικότερα τις υποψήφιες λύσεις.

- Surrogate models με Μηχανική Μάθηση: Παρά τις βελτιώσεις στη βελτιστοποίηση, οι επαλαμβανόμενες εκτιμήσεις της ομοιότητας των εικόνων για κάθε υποψήφια λύση αυξάνουν σημαντικά το υπολογιστικό κόστος. Αυτό το κόστος γίνεται μεγαλύτερο στην περίπτωσή μας λόγω του ήδη σημαντικού υπολογιστικού κόστους των intensity- based methods. Για αυτό το σκοπό, χρησιμοποιήθηκαν μέθοδοι μηχανικής μάθησης για την κατασκευή αποτελεσματικής, ακριβούς και υπολογιστικά φθηνής προσέγγισης της συνάρτησης ομοιότητας των εικόνων. Από τις μεθόδους που μελετήσαμε, η μέθοδος Support Vector Regression (SVR) ήταν η μέθοδος με το ελάχιστο σφάλμα προσέγγισης. Η μερική αντικατάσταση της αρχικής συνάρτησης ομοιότητας από την SVR μείωσε το υπολογιστικό κόστος της βελτιστοποίησης μέχρι και 47% χωρίς ταυτόχρονη μείωση της ποιότητας των αποτελεσμάτων.

- Μέτρο Ομοιότητας: Οι intensity-based methods συγκρίνουν τις εντάσεις των pixels μέσω στατιστικών ή άλλων μαθηματικών μέτρων. Το πιο επιτυχημένο και ευρύτατα χρησιμοποιημένο είναι η Αμοιβαία Πληροφορία (Mutual Information). Η επιτυχία αυτών οδηγησε στην έρευνα και άλλων παρόμοιων στατιστικών μέτρων. Στην παρούσα έρευνα δόθηκε έμφαση σε μέτρα ομοιότητας βασισμένα στη στατιστική απόκλιση του Renyi, η οποία έχει χρησιμοποιηθεί στην ταύτισης εικόνων. Στα πλαίσια της έρευνας, έγιναν πειράματα για την κατανόηση της σχέσης μεταξύ της παραμέτρου a του Renyi και του ποσοστού δειγματοληψίας της εικόνας και πως επηρεάζουν το λάθος της ταύτισης εικόνων. Τα αποτελέσματα έδειξαν ότι η αύξηση της παραμέτρου a μειώνει

το λάθος της ταύτισης ακόμα και όταν χρησιμοποιούμε μικρό ποσοστό της εικόνας.

- Συρραφή εικόνων: Τα παραπάνω επιτεύγματα χρησιμοποιήθηκαν για την κατασκευή μία νέας μεθόδου για Συρραφή εικόνων (Image Stitching), οι οποίες έχουν ληφθεί από μικροσκόπιο. Αυτή η μέθοδος συγκρίθηκε με μία μέθοδο από το Fiji, ένα πακέτο ελεύθερου λογισμικού ανοικτού κώδικα για επεξεργασία εικόνας, το οποίο χρησιμοποιείται από νευροεπιστήμονες, ιατρούς και βιολόγους. Σε αντίθεση με τη μέθοδο του Fiji, του οποίου το αποτέλεσμα εξαρτάται από εναν μεγάλο αριθμό παραμέτρων, η μέθοδός μας χρησιμοποιεί λιγότερες παραμέτρους. Αυτό καθιστά την εφαρμογή μας εύχρηστη για το μέσο χρήστη ώστε να κατασκευάσει υπερεικόνες μεγάλης ανάλυσης από το συνδυασμό μικρότερων εικόνων.

Συμπερασματικά, παρά την αύξηση των δημοσιεύσεων στην ταύτιση εικόνων με τη χρήση μεθευρετικών μεθόδων βελτιστοποίησης των intensity-based methods, τα τελευταία 15 χρόνια, εξακολουθεί να υπάρχει ένα ερευνητικό κενό στον τομέα αυτό, εξαιτίας των παρακάτω λόγων:

- Μεγάλος αριθμός των Μεθευρετικών μεθόδων βελτιστοποίησης: Υπάρχει ένας μεγάλος αριθμός μεθόδων, όπου οι περισσότερες έχουν δοκιμαστεί ελάχιστα ή καθόλου. Επιπλέον, θα ήταν χρήσιμο να ερευνηθεί η ενοποίηση μεθευρετικών και συμβατικών μεθόδων, κάτι το οποίο δεν προέκυψε από τη βιβλιογραφική μας αναζήτηση.

- Παρά το ότι τα Surrogate Models, έχουν επιτυχώς χρησιμοποιηθεί σε διάφορα προβλήματα βελτιστοποίησης, δεν έχουν χρησιμοποιηθεί στην Ταύτιση εικόνων. Η χρήση τους μπορεί να μειώσει το υπολογιστικό κόστος και τη διάρκεια της διαδικασίας.

Σε αυτήν την ερευνητική διαδικασία, ενσωματώθηκαν τεχνικές σε μεθόδους ταύτισης εικόνων με τη χρήση μεθευρετικών αλγοριρίθμων (σε μετασχηματισμούς rigid και affine) τόσο ως προς τη χρήση μιας μεθόδου βελτιστοποίησης (η οποία δεν είχε χρησιμοποιηθεί ως σήμερα) όσο και στην εισαγωγή των Surrogate Models. Παρά τις βελτιώσεις που εισήχθησαν, υπάρχει περιθώριο για επιπλέον έρευνα επεκτείνοντας την ερευνητική διαδικασία μου σε άλλα πολυπλοκότερα προβλήματα ταύτισης εικόνων και προηγμένες μεθόδους ελαχιστοποίησης κόστους για μεγαλύτερη μείωση του κόστους της ταύτισης δεδομένων χωρίς απώλεια ποιότητας των αποτελεσμάτων.

Τα αποτελέσματα της ερευνητικής δραστηριότητας έχουν μέχρι τώρα παρουσιαστεί στην επιστημονική κοινότητα μέσω των παρακάτω δημοσιευμένων εργασιών:

1. Spanakis, C., Mathioudakis, E., Kampanis, N., Tsiknakis, M., & Marias, K. (2016, October). A new approach in image registration. In 2016 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 449-453). IEEE.

2. Spanakis, C., Mathioudakis, E., Tsiknakis, M., & Marias, K. (2018, October). Elitism in intensity-based image registration. In 2018 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 1-5). IEEE.

3. Spanakis, C., Mathioudakis, E., Tsiknakis, M., Kampanis, N., & Marias, K. (2018, July). Function Approximation for Medical Image Registration. In 2018 41st International Conference on Telecommunications and Signal Processing (TSP) (pp. 1-5). IEEE.

4. Spanakis, C., Mathioudakis, E., Kampanis, N., Tsiknakis, M., & Marias, K. (2019). Machine-learning regression in evolutionary algorithms and image registration. IET Image Processing, 13(5), 843-849.

5. Spanakis, C., Mathioudakis, E., Kampanis, N., Tsiknakis, N., & Marias, K. (2019, December). Renyi divergence and non-deterministic subsampling in Rigid Image Registration. In 2019 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 1-6). IEEE.

Abstract

In image processing, it is quite common to compare images that depict the same scene or similar scenes. Often, these images are acquired from different viewpoints, at different times, under different lighting conditions or using different media. For example, in Medical Imaging, due to the complementary nature of the imaging methods, we need to compare images from different sensors (CT-MRI, CR-PET, etc.) in order to assess the condition of the patient and the treatment plan. Medical Image Registration with minimum error contributes to diagnosis that is more accurate and, therefore, better treatment and faster recovery of the patient, to his shorter stay in the hospital, with the consequent reduction of medical costs. Images from satellites are often compared in order to detect natural or man-made changes in the environment, supervise the natural resources or update maps. Even the smallest error in remote sensing Image Registration may correspond to deviation of several kilometers.

Despite its application in several areas, Image Registration remains an unsolved problem with respect to image similarity methods and optimization. The purpose of this thesis is the design of new, robust methods for the automatic alignment of image data. To this end, an approach based on different metaheuristic methods, in combination with machine learning and further search in intensity-based similarity methods is developed. Since we tried to develop a novel and robust automatic method, that can be used globally on any possible image pair. The data we used are from various applications. The thesis can be divided into three areas of focus:

- Meta-heuristic optimization: From a mathematical point of view, image registration is an optimization problem, where the objective function is the similarity (dissimilarity) of the images and the images are aligned, when this image is maximized (minimized). Therefore, optimization method must be robust and able to optimize the objective function using an few iterations as possible.

- Computational cost: In meta-heuristic optimization, the repetitive estimations of the objective function increases the computational cost. Therefore, we did research on minimizing the computational cost with the use of Surrogate models, which are constructed via machine learning methods. Although these are often used in metaheuristic optimization, they have never been used in image registration.

- Similarity measure: The similarity measure is an important key in image registration, because it is used for the comparison of the images. In our research, we focused on the intensity-based similarity measures, because of their advantages. The purpose is to find a way to minimize the effect of their disadvantages. In our work, we focused on metrics based on Renyi divergence, which have been mostly used in image registration.

Respectively, our achievements are the following ones:

- Optimization: Initially, we started with genetic algorithms which have been widely used in image registration. Progress has been made on our search regarding the relation between mutation rate and the number of the elites in elitist genetic algorithms and how they affect optimization. An increased number of elites or/and increased mutation rate leads to more robust optimization. Still, due to their inherent disadvantages, we later tried to focus on other optimization methods that have not been used in image registration. Harmony Search was a good candidate due to its simplicity and ability to exploit more sufficiently the candidate solutions. Also, it has not been used in image registration. During our research, we initially managed to modify the basic Harmony

Search optimization method into a new, better method which combines harmony search with ALOPEX algorithm. This method has been compared with ITK (Insight Toolkit for Segmentation and Registration) methods and has out competed them in terms of accuracy. Also, we devised a new termination criterion based in the combination of the mean and the standard deviation of the similarity estimation of the candidate solutions. Secondly, we expanded this algorithm and devised a new one, which has exploits even more sufficiently and faster the candidate solutions.

- Machine-Learning-based Surrogate Models: Despite the improvements in optimization, the repetitive estimations of the similarity of the images for each candidate solution increases the computational cost. This cost is even greater due to the inherent computational cost of the intensity-based methods. Here, we used machine learning methods in order to construct Surrogate Models that produce an efficient, accurate and cheap estimation of the images' similarity. Among the methods we studied, Support Vector Regression (SVR) was the method with the minimum approximation error. The partial substitution of the original similarity method by SVR reduced the minimization cost up to 47% without quality loss of our results.

- Similarity measure: Intensity-based methods compare the pixel intensities of the images via statistical or mathematical measures. The most successful and widely used is Mutual Information and its derivatives. Their success has spawned research on other similar information-based statistic measures. Our research has focused on similarity metrics based on Renyi divergence, which has been widely used in image registration. In our experiments, we tried to understand the relation between the parameter α of Renyi divergence and the sub-sampling factor and how it affects the registration error in image registration. The results have shown that the increase of the parameter α reduces the registration error even when the used image sample is small registration error even when the used image sample is small.

- Image Stitching: Last but not least, the achievements above were used in practice for the creation of a new method for stitching images acquired from microscope. This method has been compared with the one used in Fiji, an open source image processing package that has been used by neuroscientists, doctors and biologists. Unlike, the Fiji method, whose results depend on a number of parameter values inserted by the users, our method has fewer parameters which make it easier for the average user to use and create hyper-analysis images via combination of smaller ones.

Overall, we claim that, despite the increase of the publications on metaheuristic-optimization-oriented intensity-based image registration over the last fifteen years, there is still a gap in the research in that area, due to the following reasons:

- Vast number of Metaheuristic Optimization methods. There is a huge number Metaheuristic Optimization methods in existence (most of them have never been used at all or have not been exploited thoroughly), each one of them having its own advantages and disadvantages. Also, research is needed in the fusion of metaheuristics and conventional optimization methods, which, as far as we know from the literature review, is

- Despite the fact that Surrogate Models have been successfully used in other optimization problems, they have not been used in image registration. The use

of Surrogate Models can reduce the computational cost and reduce the duration of the process.

In this thesis, we added a new component in metaheuristic-oriented intensity-based rigid and affine registration regarding not only the use of a previously unused optimiza-tion method, but also the introduction of Surrogate models in image registration. Despite the improvements we introduced, there is a lot of work that can be done by expanding our work in other more complex registration problems and advanced reduction cost methods for further total reduction of the registration process without quality loss. Future research in image registration includes the combination/merge of more reduction cost methods as well as the integration of machine learning in metaheuristic optimization with the purpose of meta-optmimization.
The results of the research activity have so far been presented to the scientific community through the following published:

1. Spanakis, C., Mathioudakis, E., Kampanis, N., Tsiknakis, M., & Marias, K. (2016, October). A new approach in image registration. In 2016 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 449-453). IEEE.

2. Spanakis, C., Mathioudakis, E., Tsiknakis, M., & Marias, K. (2018, October). Elitism in intensity-based image registration. In 2018 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 1-5). IEEE.

3. Spanakis, C., Mathioudakis, E., Tsiknakis, M., Kampanis, N., & Marias, K. (2018, July). Function Approximation for Medical Image Registration. In 2018 41st International Conference on Telecommunications and Signal Processing (TSP) (pp. 1-5). IEEE.

4. Spanakis, C., Mathioudakis, E., Kampanis, N., Tsiknakis, M., & Marias, K. (2019). Machine-learning regression in evolutionary algorithms and image registration. IET Image Processing, 13(5), 843-849.

5. Spanakis, C., Mathioudakis, E., Kampanis, N., Tsiknakis, N., & Marias, K. (2019, December). Renyi divergence and non-deterministic subsampling in Rigid Image Registration. In 2019 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 1-6). IEEE.

Acknowledgments

Chapter 1

Introduction

As it is aforementioned, the purpose of this thesis is the design of new innovative and robust methods for the automatic alignment of image data. For that purpose, we describe image registration as well as the problems and the methods that have been designed.

1.1 Image Registration

Image registration [1,2] is the process of transforming image data in a single coordinate system. These data maybe multiple images, often acquired at different times, from different sensors or under different lighting conditions. In its simplest form, we have two images, known as **Floating/Source** image and **Reference/Target/** image. In this case, the first image is geometrically transformed so that their common features occupy the same place in the common coordinate system.

Image registration is an important field in image analysis and computer vision because it is widely used in a number of different areas either as a preprocessing step of more complex processes or as such.

1.1.1 Medical Imaging

Despite the technological advances in medical imaging, there is no medical imaging system that can depict in clarity the condition of the patient. Different imaging methods depict related but different aspects of the patient's condition and the information they provide the doctors with is, at best, complementray. For example, CT (Computer Tomography) depicts bone details more accurately, while MRI (Magnetic Resonance Imaging) provides the doctors with details regarding the soft tissues. Both methods lack the ability of PET (Positron Emission Tomography) and SPECT to depict accurately biological processes [3]. On the other hand, PET and SPECT delinate poorly anatomy. Therefore, it is crucial to align medical images [4] in order to obtain all the information about the patient with minimum registration error. In the area of medical image analysis, image registration has been widely used in diagnosis [5], treatment planning [4, 6, 7] and image- guided surgery [8–12]. It has also been used in more special applications such as registration of pre- and post- operative images in surgical interventions [4] and atlas mapping. Medical image registration with minimum error contributes to a more accurate diagnosis, proper medical strategy to address the health problems and, consequently, to faster treatment and recovery of the patient as well as reduction of the medical costs.

(a): Floating Image

(b): Reference Image

(c): Corrected subfigure (a)

(d): Difference Image
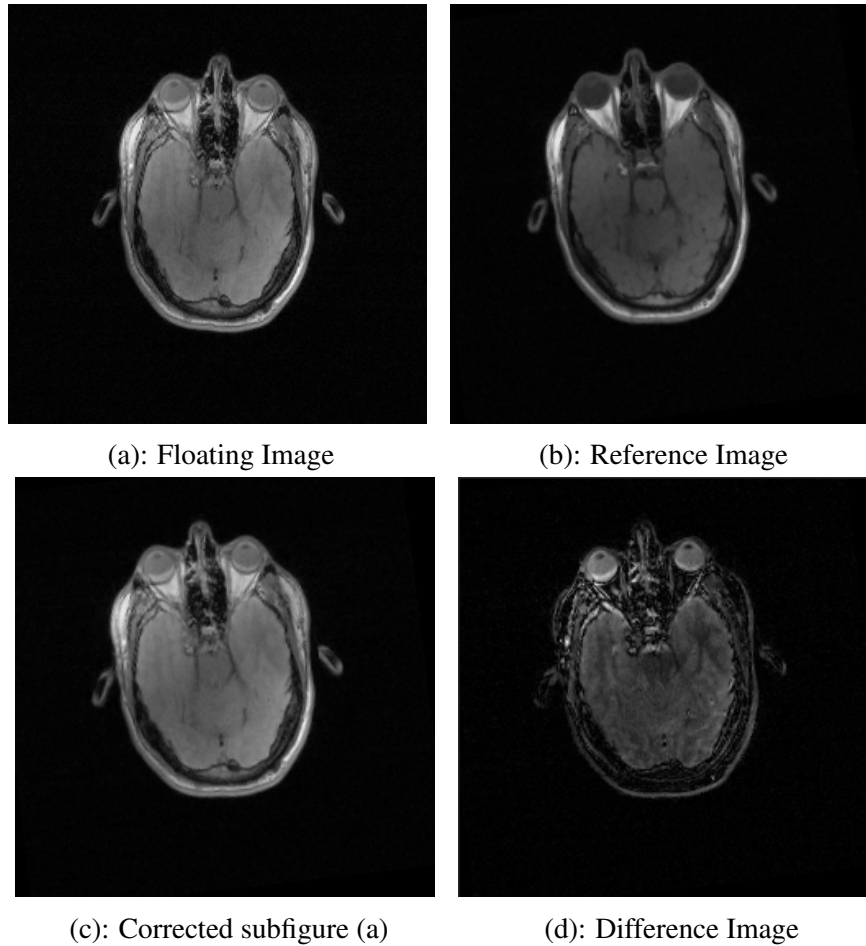
Figure 1.1: Example of Medical image Registration: (a)

1.1.2 Remote Sensing

In remote sensing, Earth and planetary scientists need to create mosaics of satellite images and observe the changes through the course of time for environmental and basic science studies [13]. More analytically, image registration in remote sensing can be classified [13] in the following way:

- Multimodal registration: Alignment of data acquired from different sensors [14]. It enables the integration of complementary information. Some of its applications are agricultural and crop forecasting, water urban planning and disease control.

- Temporal registration: It can be used for the detection of changes and surveying resources (natural and agricultural) as well as features that change over time.

- Viewpoint registration: Here we have integration of information from one moving platform or multiple platforms into 3D models. It is widely used in application such as landmark navigation [15, 16], flying formation and planet navigation [17].

- Template registration: Applications in this category such as Object search and map updating search for correspondence between new sensed data and previously developed model/dataset.

The environmental changes, whether natural or man-made, on the landscape over time makes the comparison between the images difficult. The lighting conditions,

(a): Floating Image   (b): Reference Image

(c): Corrected subfigure (a)   (d): Difference Image
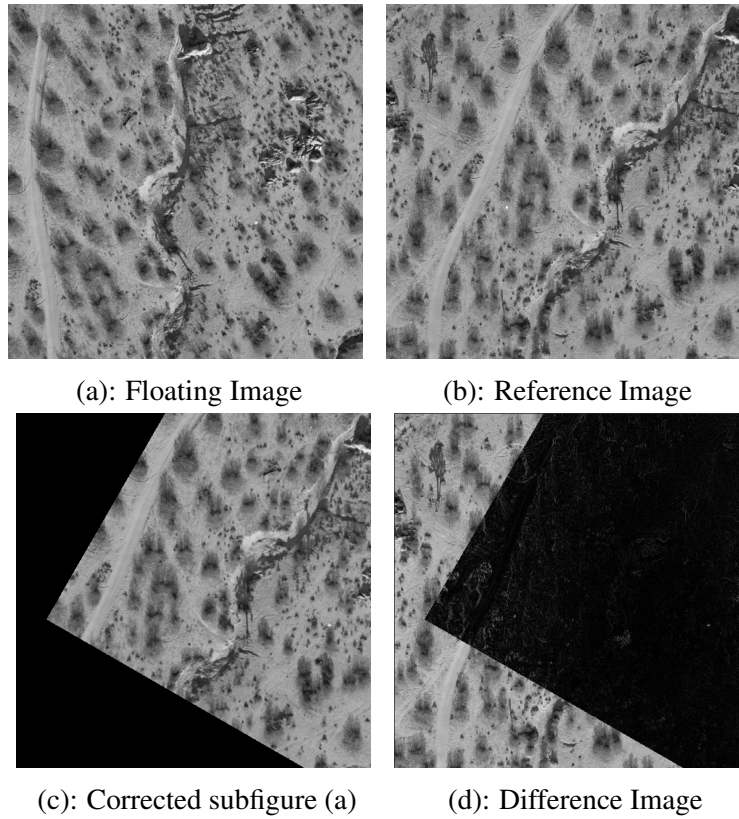
Figure 1.2: Example of Rigid image Registration: (a)

the weather patterns (clouds or fog) as well as the different image sensors make image registration a challenge.

Studies in [18, 19] have shown that even a small registration error is translated into a divergence of the range of kilometers. This can lead to the miscalculations of global measurements such as the computation of the Normalized Difference Vegetation Index (NDVI).Therefore, the image registration methods in Remote Sensing must be robust and the errors must be as small as possible.

1.1.3 Image Stitching

Image Stitching [20] is the process of creating high-resolution panoramas through the combination of multiple images with overlapping fields of view. It is widely used in many applications, some of which are the following ones:

- Medical Imaging [21]

- Video Stitching [22]

- High resolution mosaics in satellite imaging

The process of stitching is divided into the following sub-processes:

1. Image Registration: The process of aligning the images so that their mutual characteristics occupy the same space in the common coordinate system.

2. Calibration: This process aims to minimize differences between an ideal lens models and the camera-lens combination that was used, optical defects such as distortions, exposure differences between images, vignetting, camera response and chromatic aberrations.

3. Blending: It involves the adjustment of the colours in order to deal with exposure differences.

Image Registration is the first step and the "sine qua non" factor in Image Stitching, where the images are aligned in order to create the mosaic. Without correct image registration, it is impossible to create a proper mosaic that will depict the details we want to present. Therefore, since multiple images are involved, the image registration methods need to be robust and with minimal error in each image pair.

(a): First image   (b): Second Image
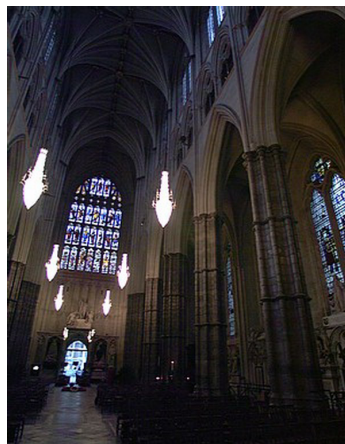
(c): Third Image

(d): Result of Image Stitching

Figure 1.3: Example of Image Stitching

1.1.4 Robotic Vision

Viewpoint calibration [23] is a method of manipulating hand-eye for the generation of calibration parameters for viewpoint control and object grasping. Apart from interaction with objects, path planning and dynamic perception are useful for successful navigation [24] in human environments. Image Registration is essential in robotic vision [25], because accurate vision sensor calibration and robust robot-vision control are the "sine qua non" factors for the development of autonomous and intelligent robotic systems [23]. Image Registration methods must be robust, accurate and fast. Although many methods use feature-based comparison methods for faster registration [23–25], intensity-based methods [26] have started to gain popularity.

1.2 Research Methodology for Image Registration

In image registration, the main unresolved needs are the following ones [27]:

- Accuracy: It is an important factor for the performance of image registration. For example, in remote sensing, an image registration error by only one pixel can introduce more than 50% error in change detection applications via use of Normalized Difference Vegetation Index [28]. Another example is Augmented Reality Navigation in dental surgery. In such procedures, highly precise operations are required, where the accuracy error must be less than 1mm. [29].

- Robustness: Image Registration can be modeled as an optimization problem, where the objective function is an image similarity function whose arguments are the transformation. The alignment of considered successful, if the global optimum of the objective function is found, i.e. the transformation that maximizes the similarity of the images. The optimization method should be able to cope with all image pairs regardless of their origin, lighting conditions and the timepoints at which they are acquired.

- Duration/Computational Cost: In many cases, especially where the objective function has many Degrees of Freedom, extensive search must be applied for the location of the global optimum. This can easily increase the computational cost/duration of the image registration process, especially if the objective function is costly. In this case, it is a challenge to create an real-time image registration application. Such an applicaion can be extremely helpful in cases such as real time surgery [29].

- Similarity measures: Similarity measures is an important criterion to evaluate the similarity of the images (for a given transformation) during the image registration process. A similarity measure must be robust and ab,e to describe quantitatively the similarity of two images for any given transformation. The most common way to compare images is to use either common features or statistic measures. The latter has been applied successfully in multi-modal image registration where it is more difficult to find common features between the images.

- Convergence to local optima: In the context of image registration, optimization is the location of the optimal transformation that best aligns the images. Many optimization techniques have been employed for image registration such as Conventional optimization methods (e.g. Powell's method and Levenberg-Marquardt alorithm) and Metaheurstic methods.

- Automatic image registration: It is very helpful in cases where there is need for minimum human interaction such as image-guided surgeries [27]. They can be efficient as they require less time and minimum efforts from user during the image registration process. However, their performance depends highly on accuracy and optimization algorithms.

The rationale behind the research presented in this thesis is to develop new automatic image registration methods that can cover the above needs as much as possible. Being able to align images with maximum accuracy and robustness and minimum Duration/Computational cost can be useful in many areas where it is used. Therefore, it is important to study the parameters that contribute to whether these needs are met, how they have modelled the numerous image registration methods and what more is presented in this thesis. In the next chapters, a detailed analysis for the research will be presented. More analytically, in the second chapter a literature review will be presented regarding these factors as well as the recent trends in intensity-based image registration and how they address its problems. In the third chapter, the methodologies as well as the methods that were used for the improvements in image registration. The results will be discussed in the fourth chapter. Finally, in the fifth chapter the conclusion and the future plans will be presented.

Chapter 2

Mathemacical Methods

2.1 The problem of Image Registration

The issues of accuracy, robustness, computational cost/duration, automation and similarity measures and local optima convergence are determined by the following parameters:

- Deformation modeling

- Similarity estimation

- Optimization methods

Each one of these parameters affects the performance of an image registration method not only directly, but also indirectly by contributing the severity of the others. For example, an increase of the complexity of the deformation model leads to an increase of the Degrees Of Freedom (DOF) of the optimization problem. The increase of DOF increases exponentially the search space, which requires a robust optimization method to solve the problem.
On the other hand, the similarity method may have several local optima (regardless of the complexity of the deformation model). Also the similarity method must be robust and salient, regardless the lighting conditions, sensors or different times, at which the images were acquired.
Optimization methods must be able to find the global optimum at the minimum computational cost, no matter the images we want to align. Out of these open parameters, optimization methods have been the most important one, upon which extensive reasearch has been conducted over the past twenty years, producing a number of methods that deal with that problem. The ability of the optimization method to locate the global optimum, regardless of the degree of freedom of our problem is an open problem in image registration, due to the fact that there are many optimization methods (each one with its own advantages and disadvantages). Also, due to the fact that there is no single method that effectively compares any image pair to be aligned, extensive research has been done on intensity-based methods [30] due to their ability to align images from different modalities, which makes it difficult to find common features between them. What we propose, are the following:

- Use of simpler optimization methods with minimum dependence from parameters.

- Use of machine learning as a means of computational cost reduction.

- Study of other similarity methods.

2.2 Current methods and their Main Drawbacks

Due to the vast range of applications to which image registration is applied, it is not surprising that many methods have been developed that address the problem from a different perspective. However, there is not a general method that is optimized for all uses. To be more exact, image registration methods differ from each other with respect to the different approaches [31] of:

- Deformation Models: In image registration, we need to know the transformations that need to be applied on the **Source** image so that it can be properly transformed and, thus, be aligned to the **Target** image.

- Similarity/Comparison: An important problem is how to calculate their similarity. Some methods use image features, while other methods use pixel values for the comparison of the images.

- In Image Registration, the goal is to find the transformation that maximizes the similarity between two images. From a mathematical point view, it is an optimization problem where the objective function is the similarity between two images, whose estimation depends on the transformation we apply on the image.

2.2.1 Deformation Models

The choice of the deformation is important as it defines, on one hand, the richness of description of the transformation and, on the other hand, the computational efficiency. In Figures 2.1-2.4, examples of transformations are presented, where, in each subfigure, the left image is the original image and the right image is the result of the respective transformation. The number of parameters that the image registration process estimates throughout the deformation range from 3 in the case of rigid registration (6 if the images are 3D) to millions in the case of non-parametric dense transformations.

Rigid Transformation is described by Eq.2.2, where x is the rotation angle and t_x, t_y is the translation along axis x and y respectively. It is mostly used in cases where there is no affine transformation (such as shearing and scaling) or elastic ones. One example of rigid registration is image registration of brain images [32–38] or as pre-processing step for non-rigid registration [39]. The greatest challenge is the elastic or non-linear transformations. Unlike linear transformations, the non-linear ones tend to be local and non uniform. The problem in non-rigid registration is the mathematical modeling of the deformation [40].

$$T_{Translation} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{pmatrix} \tag{2.1}$$

$$T_{Rigid} = \begin{pmatrix} \cos(x) & \sin(x) & 0 \\ -\sin(x) & \cos(x) & 0 \\ t_x & t_y & 1 \end{pmatrix} \tag{2.2}$$

$$T_{Affine} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix} \tag{2.3}$$
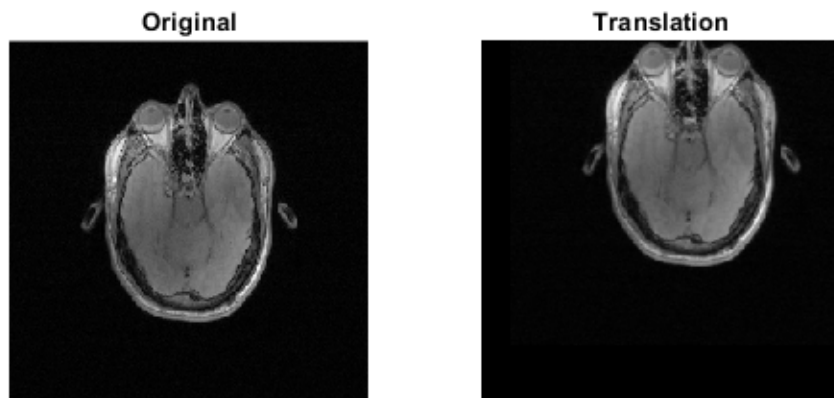
Figure 2.1: Translation example
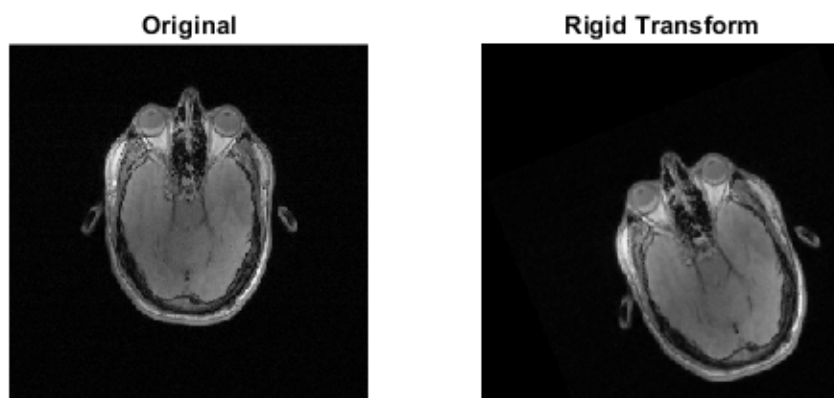
Figure 2.2: Rigid transformation example

2.2.2 Similarity methods

Similarity methods assess the physical correspondence between the common features of the images that are compared. The ideal image comparison metric should be able to compare images irrespective of the sensors (same or different), lighting conditions
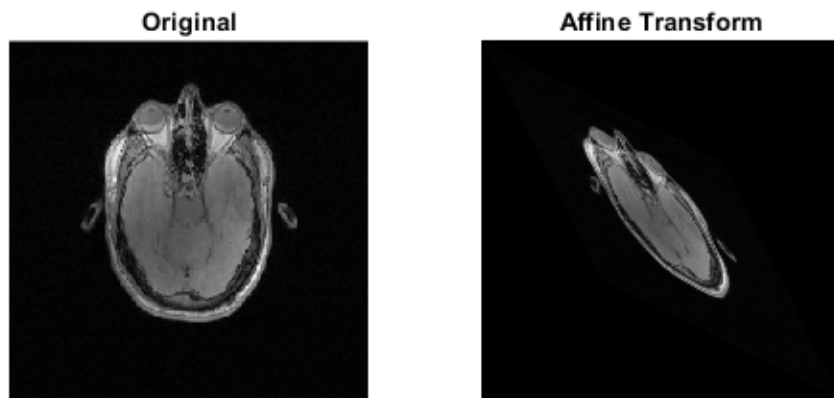
Figure 2.3: Affine transformation example
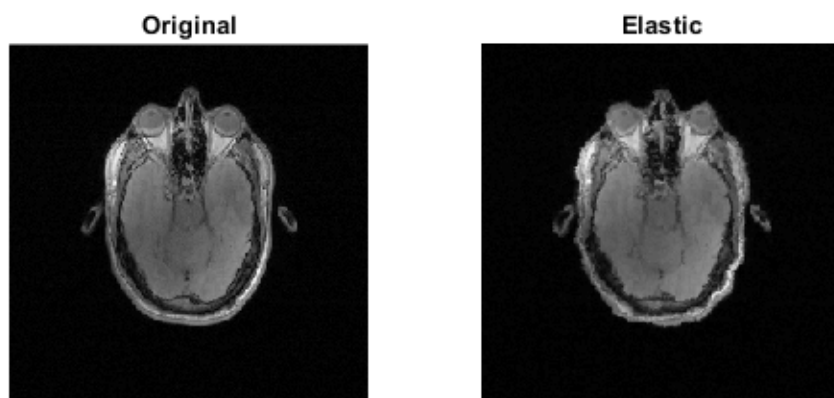
Figure 2.4: Elastic transformation example

and the difference in the time of the acquirement.

Most similarity methods are classified into the following categories [41]:

- Feature-based methods: These methods use the common features of the images such points [42] (SIFT [43] and SURF [44, 45]), edges [46–48] and contours

[49] for the comparison of the images. Since, they use a small part of the images for the similarity estimation, they are fast. They are mostly used in cases where the images are rich in details such as remote sensing and computer vision. On the other hand, in medical imaging the images are not rich in such details. Also, it is difficult to find distinctive and salient features that can be used in different lighting conditions and multi-modal imaging. This can be partially be dealt with via manually placed landmarks. Last but not least, there may be in general greater need for preprocessing.

- Intensity-based methods: Unlike the methods of the previous category, intensity-based methods compare the pixel patterns of images via mathematical or statistical measure. The idea is that two images are perfectly aligned, when the measure is maximized/minimized. Some of the most common intensity-based methods are Mean Squared Difference (MSD), Cross-Correlation, Mutual Information and Normalized Mutual Information. The advantage of these methods lies in their robustness and the need for minimal, if any, preprocessing of the images.Some of them are useful for monomodal registration such as sum of mean. Mutual Information and Normalized Mutual Information have been successfully used in multimodal registration producing robust and reliable results. In fact, their wide and successful use has spawned a family of new information-based statistic measures such as Renyi divergence, Tsalis divergence. However, their robustness and accuracy can be affected [41] by the way of estimating the probability distributions and the choice of the interpolator.
Another disadvantage is the need to use all the pixel intensities (or at least a significant percetage of them) for the estimation of the similarities. This can lead to an increase of the computational cost. Also, they rely on the assumption of independence and stationarity of the intensities from pixel to pixel [4], without taking into account their spatial dependencies. Further, the intensity relationship is assumed to be spatially stationary. As a result, such measures tend to fail when registering two images corrupted by spatially-varying intensity distortion.

- Hybrid methods: Instead of focusing exclusively on either features or pixel intensities, several hybrid methods have been constructed in order to [50–52]. In this way, it is possible to combine the advantages of both methods, while minimizing the effects of their respective disadvantages.

2.2.3 Optimization methods

Image registration is perceived as an optimization problem where we seek for the transformation that maximizes the similarity of the images. Therefore, depending on the complexity of the problem, the proper optimization method must be chosen.
The optimization methods that are used in image registration are classified into the following categories.

- Conventional Mathematical methods: Common optimization methods that are used in this category are the Powell's minimization method [33, 53], Levenberg-Marquardt [51, 54], Steepest Gradient Descent, Quasi-Newton methods [55]. These methods have been used extensively in image registration. Despite their simplicity, their success depends on the initial conditions (i.e. the initial point or/ and direction vectors). In the best scenario, good initial conditions lead to quick convergence to the global optimum. The worst

case scenario is a poor choice which leads to convergence to a local optimum, from which it is difficult, if not impossible, to escape. This problem becomes more pronounced as the number of dimensions increases.

- Stochastic methods: The inability of the conventional mathematical methods to deal with the entrapment at the areas of local optima has lead to the increasing use of stochastic optimization methods. These methods introduce stochasticity for the creation of candidate solutions to our problem, which facilitates the escape from local optima. The most common ones that have been used are stochastic annealing [56, 57], evolutionary optimization methods (Genetic algorithms [58, 59], Evolution strategy [60, 61], Differential Evolution [62, 63])) and swarm optimization methods (Particle Swarm Optimization [64–67] , Artificial Bee Colony [68]). In order to escape local optima and find the global one, they have to do repetitive estimations of the similarity of the images for all the candidate solutions they generate. This can increase the computational cost, especially in the case where even a single estimation alone may be computationally expensive.

2.3   Optimization Methods in Image Registration

When two images (called **Fixed**/**Target** and **Moving**/ **Source** respectively) are aligned (given a certain transformation), their similarity (disimilarity) is maximized (minimized) for that transformation. Therefore, the image registration problem can be transformed into an optimization problem where the objective function (i.e. the function we seek to optimize) is the similarity (disimilarity) function, whose independent variables are the transformation of the **Moving** Image. In intensity-based image registration, the objective function is a mathematical/statistical measure [41] which uses the pixel intensity values of the images, in order to compare them. Several optimization techniques have been proposed for image registration

2.3.1   Conventional Optimization Methods

Conventional Mathematical methods [55] have been the first optimization methods that have been used in image registration. Below we present some of the most well known ones.

Steepest Gradient Descent

Steepest Gradient Descent [69] is variance of Gradient Descent (1). Using the later, given an objective function $f : \mathbb{R}^N \to \mathbb{R}$, we start from an initial point $x_0 \in \mathbb{R}^N$ and we calculate a series of points x_i(Eq.2.4):

$$x_{i+1} = x_i - \gamma \nabla \left. f(x) \right|_{x=x_i} , \ i = 1, 2, ... \tag{2.4}$$

where $\gamma > 0$ is the step size.
In Table 1, the basic Gradient Descent algorithm is presented, where γ is the step size, Max_iter is the maximum number of iterations and ϵ is the tolerance. Despite its simplicity, it has two major disadvantages:

- It requires the computation of the gradient of the objective function, which can be computationally expensive, especially when the search space is large and the objective function alone is computationally expesive.

```
   Data: $\gamma$, $Max\_iter$
   Result: best individual
1  $x_0$=random_initialization();
2  for $i = 1, 2, \ldots, Max\_iter$ do
3  │   $x_{i+1} = x_i - \gamma \nabla f(x)|_{x=x_i}$
4  end
5  return best individual
```

Algorithm 1: Basic Gradient Descent

- The γ parameter affects the performance of the algorithm. A small value of γ leads to slow convergence, while large values may lead to no convergence at all. Depending on the optimization problem as well as the different phases of the optimization process, gamma needs to be adjusted.

- It gets easily entrapped in local minima.

Steepest Gradient Descent can overcome the disadvantage of the γ parameter by incorporating the gradient in order to reach the optimum.

```
   Data: $\gamma$, $Max\_iter$
   Result: best individual
1  $x_0$=random_initialization();
2  for $i = 1, 2, \ldots, N$ do
3  │   Choose $\gamma$ via backtracking or linesearch;
4  │   $x_{i+1} = x_i - \gamma \nabla f(x)|_{x=x_i}$
5  end
6  return best individual
```

Algorithm 2: Basic Steepest Gradient Descent

The choice of γ parameter may need to change during the optimization process. For example, if the initial point x_0 is far from the optimum, the value of the parameter should increase in order to facilitate the quicker convergence, while near the global optimum the value should be small to avoid steps that will lead to escape from the area of the global optimum. Therefore, in Steepest Gradient Descent, it is calculated via backtracking or linesearch using algorithms 1-D optimization algorithms such as Brent's minimization algorithm [69]. Although, more efficient than the basic Gradient Descent, it has its own distinct disadvantages:

- The linesearch along 1-D may increase the computational cost.

- As, we get closer to the optimum, the procedure is reduced to a series of tiny steps which may slows down the convergence.

Conjugate-gradient

Conjugate-gradient [70] is a method similar to Steepest Gradient Descent. However, unlike Steepest Descent, this method updates the step size γ not according to the direction of the gradient, but conjugate to the previous direction.

The calculation of β_{k-1} varies from variant to variant. For example, in the Fletcher-Reeves variant [71] $\beta_i = \dfrac{\|\nabla f(x_{i+1})\|_2^2}{\|\nabla f(x_i)\|_2^2}$, while the Polak-Ribiere variant [72] defines $\beta_i = \dfrac{(\nabla f(x_{i+1}) - \nabla f(x_i))^T \nabla f(x_{i+1})}{\|\nabla f(x_i)\|_2^2}$. When the objective function is nonlinear,

```
   Data: γ, Max_iter
   Result: best individual
1  x_0=random_initialization();
2  d_0=random_initialization();
3  for i = 1, 2, ..., N do
4      Choose γ via linesearch;
5      x_i = x_{i-1} + γd_{i-1};
6      Calculate β_{i-1};
7      d_i = -∇f(x_i) + β_{i-1}d_{i-1}
8  end
9  return best individual
```

Algorithm 3: Basic Steepest Gradient Descent

then the later invariant surpasses the previous one in terms of convergence speed [55]. Other variants are:

- Hestenes-Stiefel [73]

- Dai–Yuan [74]

Powell

Powell's direction set method [53] requires only the evaluation the objective function itself and not its derivatives. However, unlike the previous methods that require an initial point, this method also requires a set of initial search vectors. For an objective function $f : \mathbb{R}^N \to \mathbb{R}$, N search vectors $s_1, s_2, \ldots, s_N$ are passed .
Then, the function is minimized along each search vector via bi-directional search such as Golden-section Search [75] or Brent's method [76]. The algorithm is presented in Algorithm 2.

```
   Data: Max_iter,{s_i}_{i=1}^{N} vector set
   Result: best individual
1  x_0=random_initialization();
2  Create {s_i}_{i=1}^{N} random search vector;
3  for i = 1, 2, ..., Max_iter do
4      Find the minima via bi-directional line search
          {x_{i-1} + Σ_{k=1}^{j} a_k s_k}_{j=1,...,N};
5      x_i = x_{i-1} + Σ_{k=1}^{N} a_k s_k;
6      Remove from {s_i}_{i=1}^{N} the vector that has contributed most and replave it
          with Σ_{k=1}^{N} a_k s_k
7  end
8  return best individual
```

Algorithm 4: Powell's Method

The advantage of this method is that it does not need any derivatives of the function and its complexity is dependent entirely on the bi-directional method. However, a good choice of an initial point and a initial search vector set is very difficult and often leads to failure quite easily.

Quasi-Newton

Newton's minimization method (Alg. 5) uses second derivative or (in the case of multivariate optimization) the Hessian matrix ($\nabla^2(f(x_{i-1}))$).

Data: γ, Max_iter
Result: best individual
1 x_0=random_initialization();
2 **for** $i = 1, 2, \ldots, Max\_iter$ **do**
3 $\quad$ Choose γ via linesearch;
4 $\quad$ $x_i = x_{i-1} + \gamma\nabla^2(f(x_{i-1}))\nabla f(x_{i-1})$;
5 **end**
6 return best individual

Algorithm 5: Newton's minimization method

Unfortunately, this matrix can be either too expensive or even impossible to be defined. Therefore, Quasi-Newton methods are used in the case of multivariate optimization, where the aforementioned problem is solved by building an approximation of it. The general algorithm of the Quasi-Newton optimization method is described in Alg. 7.

Data: γ, Max_iter
Result: best individual
1 x_0=random_initialization();
2 $B_0 = I$;
3 **for** $i = 1, 2, \ldots, Max\_iter$ **do**
4 $\quad$ Solve $\Delta x_{i-1} = -\gamma_k B_{i-1}^{-1}\nabla f(x_{i-1})$;
5 $\quad$ $x_i = x_{i-1} + \Delta x_{i-1}$;
6 $\quad$ $y_{i-1} = \nabla f(x_i) - \nabla f(x_{i-1})$;
7 $\quad$ Compute B_i from B_{i-1}, y_k and Δx_{i-1} ;
8 **end**
9 return best individual

Algorithm 6: Newton's minimization method

---

where x_0 is the initial point, B_i, $i = 0, 1, \cdots$ is the approximation of the Hessian matrix (B_0 is the identity matrix) and γ_k is the parameter that satisfies the Wolfe conditions. The methods of calculating the B_i approximation of the Hessian matrix are presented in Table 2.1.

Table 2.1: Methods for calculating the B_i approximation of the Hessian matrix

Method	B_i
BFGS	$B_{i-1} + \dfrac{y_{i-1}y_{i-1}^T}{y_{i-1}^T \Delta f(x_{i-1})} - \dfrac{B_{i-1}\Delta x_{i-1}(B_{i-1}\Delta x_{i-1})^T}{\Delta x_{i-1}^T B_{i-1}\Delta x_{i-1}}$
Broyden	$B_{i-1} + \dfrac{y_{i-1} - B_{i-1}\Delta x_{i-1}}{\Delta x_{i-1}^T \Delta x_{i-1}}\Delta x_{i-1}^T$
DFP	$\left(I - \dfrac{y_{i-1}\Delta f(x_{i-1})}{y_{i-1}^T \Delta f(x_{i-1})}\right) B_{i-1} \left(I - \dfrac{\Delta f(x_{i-1})y_{i-1}^T}{y_{i-1}^T \Delta f(x_{i-1})}\right) + \dfrac{y_{i-1}y_{i-1}^T}{y_{i-1}^T \Delta f(x_{i-1})}$
Broyden family	$(1 - \phi)B_i^{BFGS} + \phi B_i^{DFP}, \ \phi \in [0, 1]$
SR1	$B_{i-1} + \dfrac{(y_{i-1} - B_{i-1}\Delta x_{i-1})(y_{i-1} - B_{i-1}\Delta x_{i-1})^T}{(y_{i-1} - B_{i-1}\Delta x_{i-1})^T \Delta x_{i-1}}$

Levenberg–Marquardt

Steepest descent and Newton's Optimization method are complementary in the advantages they provide. Steepest descent near the global optimum starts doing tiny steps near the global optimum. On the other hand, Quasi-Newton methods may converge rapidly, however they are particularly sensitive to the starting location. The Levenberg–Marquardt method fuses the two methods into one as we see in (2.5).

$$x_{i+1} = x_i - (H + \lambda I)^{-1} \nabla f(x_i) \tag{2.5}$$

where, H is the Hessian matrix of f, λ is the step size and I is the identity matrix.

Data: γ, Max_iter
Result: best individual
1 x_0=random_initialization();
2 $B_0 = I$;
3 **for** $i = 1, 2, \ldots, Max\_iter$ **do**
4 $\quad$ Solve $\Delta x_{i-1} = -\gamma_k B_{i-1}^{-1} \nabla f(x_{i-1})$;
5 $\quad$ $x_i = x_{i-1} + \Delta x_{i-1}$;
6 $\quad$ $y_{i-1} = \nabla f(x_i) - \nabla f(x_{i-1})$;
7 $\quad$ Compute B_i from B_{i-1}, y_k and Δx_{i-1} ;
8 **end**
9 return best individual

Algorithm 7: Quasi Newton's minimization method

2.3.2 Metaheuristic Optimization Methods

Conventional optimization methods share a common disadvantage, which is the dependence on the nature of the objective function. A convex function without local optima is an easy task. As the number of the local optima increases, it becomes easier for these methods to get trapped into areas of local optima, from which it is difficult to escape. Unless the initial conditions are not good enough, then the global optimum may not be found. This problem becomes more pronounced as the number of the function arguments increases. Especially, in the case of image registration, each image pair that needs to be aligned has a objective function used as a comparison measure that is unique, even when compared with other images pairs that are acquired from the same sensors and under the same conditions. Initial parameters that are good for locating the optimal transform in one pair are most probably useless in the case of another one.

This problem can be solved by the introduction of stochasticity in our search, i.e. introduction of elements such as gaussian noise that can help the method escape the local optimum. Research in stochastic optimization has lead to the design of optimization algorithms that are able to overcome the entrapment in areas of local optima. The most well- known categories of them are the following ones:

- Evolutionary Algorithms

- Swarm Optimization Methods

2.4 Evolutionary Algorithms in Image Registration

Evolutionary algorithms are metaheuristic algorithms whose mechanisms are inspired by the mechanisms of the Darwinian process of evolution such as reproduc-

tion, mutation and selection. Starting from an initial random population of candidate solutions, some of them are selected according to their fitness (i.e. their function estimation) in order to create new solutions. This process is repeated until the termination criterion(a) is(are) met. Below we present a brief explanation of the evolutionary algorithms along with the presentation of examplary papers where their use is mentioned.

2.4.1 Genetic Algorithms

Genetic Algorithms are one of the most well-known evolutionary algorithms that have been used in image registration. They have been used in both in rigid [77, 78] and affine [79–81] registration and in many applications such as image stitching [82] and medical imaging [83]. These algorithms are based on the Darwinian idea of evolution, according to which the fittest individuals of a group are most likely to reproduce and eventually pass their genes to the next generation, thus supplanting the least fit ones. As in evolution, new individuals are created via recombination of the genes from the selected ones of the current generation and they are then mutated. The Basic Genetic Algorithms(8) initially create a population of random candidate solutions of size $\mathbf{P}$. Then, some of them are selected based on their fitness (function evaluation) for reproduction. The fittest are most probable for selection, although the least fit ones are not totally disregarded. The probability of reproduction is set by the Cross-over Rate ($\mathbf{CR}$). After the creation of the new population, its members are mutated with probability ($\mathbf{MR}$). The mutated population replaces the current one. This process of Selection, Reproduction and Mutation is repeated until the termination criteria are met (e.g. Maximum number of iterations).

Data: Population Size $\mathbf{P}$, Mutation rate ($\mathbf{MR}$), Cross-over rate ($\mathbf{CR}$)
Result: best individual
1 Population=initialization($\mathbf{P}$);
2 **while** *termination criteria not met* **do**
3     $\mathbf{S}$=Select();
4   new_Population = Recombination($\mathbf{CR}$,$\mathbf{S}$);
5   Mutation(new_Population, $\mathbf{MR}$);
6   Population=new_Population;
7 **end**
8 return best individual

Algorithm 8: Basic Genetic Algorithm

The parameters that affect the performance of the genetic algorithms are the following ones:

- Selection

- Cross-over Rate

- Mutation Rate

Depending on the selection method as well as the values of the last two parameters, the convergence can be greatly affected. Due to the limitations of genetic algorithms, several modifications have been proposed in order to overcome them in the context of image registration. Below we present an indicative table of the GA-based approaches in image registration.

Table 2.2: Indicative GA-based approaches in Image Registration

Reference	Transformation	Description
[84]	Rigid+Elastic	In the first step of Rigid Registration, a repartition constraint on the population over the search space based on Latin Squares is introduced, thus ensuring the diversity of the population from the begining to the end of the algorithm.
[85]	Rigid+Scaling	Use of Elitist strategy with adaptive mutation rate for coarse registration and Least Square Matching for fine tuning
[79]	Affine	Use of Elitist strategy. 20% of the best solutions are kept for the next generation. 3% is produced via random mutations and the rest via cross-over along with Local gradient correction of the replication pool.
[86]	Rigid	Use of Elitist strategy. 20% of the best solutions are kept for the next generation. 3% is produced via random mutations and the rest via cross-over.
[87]	Affine	Hybridization of GA with Powell's minimization method for accurate results.
[88]	Rigid	Hybridization of GA with DIRECT, a deterministic optimization algorithm for fine tuning
[80]	Affine	The first operation is a quantum interference which allows a shift of each qubit in the direction of the corresponding bit value in the best solution
[89]	Rigid	Dynamic mutation range, starting from a large value and decrease to a smaller value as the global optimum is reached. Upon termination, a second search for the global optimum is applied at a reduced space search around the area of the best solution that is found at termination for ensuring that the finely tuned solution is as close as possible to the global optimum.
[90]	Rigid	Niche-oriented genetic algorithms, i.e. genetic algorithms where the population is divided into sub-populations, each one of them explores a different portion of the search space, retaining in this way population diversity
[91]	Rigid+Scale	Multi-resolution Approach
[57]	Rigid+Scale	Use of Elitist GA along with Neighborhood Search and Adaptive mutation rate.
[92]	Rigid	Adaptive crossover and mutation rate according to the fitness of an idividual solution. When there is need for faster convergence crossover rate increases and mutation rate decreases, while the reverse process is applied when there is need to expand search and prevent stagnation.

[93]	Affine	Niche-oriented genetic algorithms, i.e. genetic algorithms where the population is divided into sub-populations, each one of them explores a different portion of the search space, retaining in this way population diversity and minimizing the probability of entrapment to local optima.
[94]	Affine	Multi-resolution Approach
[95]	Projective	The algorithm starts three times from the beginning, preserving only the best individual from the previous step and re-initialising the whole population.
[96]	Translation	Genetic-algorithm-based image registration is used as a preliminary step of Image Stitching
[97]	Rigid	This algorithm combines GA with Simulated Annealing for Local Search.
[77]	Rigid	One-point crossover is performed. Mutation of each bit in the chromosome string of the offspring individuals is allowed to change value with a small mutation probability. Elitism and Fitness sharing are used for faster convergence and population-diversity maintanance respectively.
[98]	Affine	Elitist GA allows two individuals to be kept unchanged and be integrated into the next generation.
[99]	Rigid+Scaling	At each iteration, from the current population of size P, the m highest solutions are copied directly to the next generation and the P-m individuals are Selected via roulette wheel selection. The result of the genetic algorithm will be the initial point for Powell's optimization method.
[100]	Affine	Multi-resolution Approach
[101]	Rigid	Adaptive crossover and mutation rate in order to avoid local optima. Employment of second crossover and migration strategy in order to avoid stagnation.
[102]	Rigid	Use of Blend Crossover and random mutation, as well as restart strategy
[83]	Rigid	Multiresolution along with elitist genetic algorithm, plus Downhill Simplex for fine tuning.

[78]	Affine	The approach here has two innovations:
		• The first resolution is performed a fixed number of times, independently of its outcome. At the end of this process, the best solutions found are considered for the second resolution. As the first resolution deals with a low-resolution version of the input images, this stage of the registration is cheap in terms of the total computational effort.
		• In the second resolution, which is the phase of refinement, the searcg space is altered according to the best solution. Instead of having, for each transformation parameter, the original range $[l, u]$ is replaced by $[b - (b - l)/h, b + (u - b)/h]$, where b is the value of the parameter in the best solution and h is the shrinking factor.
[82]	Projective	In selecting initial population, GA chooses N initial individuals randomly in feasible interval. After the introduction of chaos, it selects L (L is the number of model parameters) initial values, generating L chaos sequences after n (n >N) iterations. These chaotic variables are added to parameter's intervals and there will be n feasible solutions. Select N better solutions as the initial population by calculating their fitness functions. Also, in each loop, a small-scale disturbance is added to the better individual and it is iterated for many times. Choose the best one to keep in the population. This will do well to the optimization and prevent the local extrema.
[103]	Affine	This variant used blend crossover (BLX-α) and random mutation. An important adaptation is the use of multiple resolution along with restart and a search space adaptation mechanism. Last, but not least, since the second resolution is the a refinement phase, the search is focused around the best solution via the range restriction of the transformation parameters.
[104]	Rigid+Scale	A new addition is the use of arithmetic crossover, where the two offsprings q and p instead of inheriting variable values from parent a and b from the second one, they are assigned the following values: $p = \gamma a + (1 - \gamma)b$, $q = (1 - \gamma)a + \gamma b$, $\gamma \in [0, 1]$

[105]	Translation	Micro Genetic Algorithm (μGA) is a variant of GA, distinguished by the use of high Crossover rates and, most importantly, the use of re-initialization in lieu of mutation.

2.4.2 Evolution Strategies

Evolution Strategies (ES) [106, 107] is a family of evolutionary algorithms, which unlike the genetic algorithms, relies heavily on mutation and selection for the search of the global optimum. The simplest form is the $(1 + 1)$ ES. Starting from an initial random candidate solution, we create a new candidate solution by randomly adding a normally distributed random value to each of the elements of the current solution. If the new is better than the current one, then the first replaces the second. Otherwise, the new one is discarded. This process is then repeated until the termination criteria are met. The algorithm is the following one:

Data: Step size
Result: best individual
1 Create initial candidate(**P**);
2 **while** *termination criteria not met* **do**
3   **newP = P**+GaussianStep(step);
4   **if** *newP is better than P* **then**
5   P=newP;
6   **end**
7 **end**
8 return best individual;

Algorithm 9: Basic (1+1)ES

The most important factor is the step size of the search. Very small steps lead to an unecessary large number of iterations, while a large size step may lead to a crudely approach the optimum and then moving far away from it. Therfore, instead of a predetermined step size, it is often to use adaptable step size in our search for the global optimum. A rule regarding the step size that has been widely used is the $1/5$ success rule. According to that rule, after every n mutations, the number of successful mutations is checked. If this number is less than $n/5$, increase the step, else if it is more than $n/5$, reduce the step.

ES(μ, λ)

Unfortunately, in the previous algorithm, only one candidate solution is updated and, therefore, tasked to find the global optimum. This intensifies the dependence on the step length. In the case of a large search space, the convergence to the global optimum is slow.

In order to solve this problem, an enhanced version of the previous algorithm is used, named ES(μ, λ). Initially, random population of μ candidate solutions is created. In each iteration, each one of the population produces λ/μ offsprings on average, so that the total number of them is λ. The creation of the offspring is applied with the same way as in ES(1+1). Then, the best μ of the λ offspring replace the current population. An elitist variant of this algorithm is ES($\mu + \lambda$), which in each iteration combines the current μ-size population with the λ-size population into a single larger population, from which the best μ candidate solutions are used to replace the μ-size

population. The reason that makes it an elitist version is that if some of the elements of the current population are better than the new candidate solutions, they are not replaced as in ES(μ, λ), but are kept and integrated into the new population. In the next table, present indicative research papers where ES has been used for image registration [108].

Table 2.3: ES-based Image Registration methods

Reference	Transformation	Description
[109]	Affine	Combination of ($\mu + \lambda$)ES with Multiscale
[110]	Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[111]	Rigid+Nonrigid	Covariance Matrix Adaptation (CMA) evolution strategy+Multiresolution Strategy
[61]	Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[112]	Nonrigid	Covariance Matrix Adaptation (CMA) evolution strategy
[113]	Projective	Covariance Matrix Adaptation (CMA) evolution strategy
[114]	Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[115]	Rigid+Nonrigid	Covariance Matrix Adaptation (CMA) evolution strategy (μ, λ)
[116]	Rigid	Multi-resolution Covariance Matrix Adaptation (CMA) evolution strategy
[117]	Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[118]	Piecewise-Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[119]	Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[120]		Covariance Matrix Adaptation (CMA) evolution strategy
[121]	Affine	Covariance Matrix Adaptation (CMA) evolution strategy
[122]	Rigid	Multi-Start Covariance Matrix Adaptation (CMA) evolution strategy
[123]	Rigid	Covariance Matrix Adaptation (CMA) evolution strategy
[124]	Rigid+Scale	Hybridization of ES with PSO

2.4.3 Differential Evolution

Differential Evolution (DE) [125] is a method that optimizes a problem via combination of the positions of existing solutions from the population. If the new position of a solution is an improvement, then it is accepted and is inserted into the population, otherwise the new position is simply discarded. In 12, the basic Differential Evolution algorithm is presented.

```
    Data: Step size
    Result: best individual
1  Create initial candidate(P);
2  while termination criteria not met do
3  │   newP = P+GaussianStep(step);
4  │   for i=1:μ do
5  │   │   Produce ⌊λ/μ⌋ offisprings based on the i−th parrent;
6  │   end
7  │   Replace the parrent with the best μ offsprings
8  end
9  best individual=best of the current population return best individual;
```

Algorithm 10: Basic ES(μ, λ)

```
    Data: Step size
    Result: best individual
1  Create initial candidate(P);
2  while termination criteria not met do
3  │   newP = P+GaussianStep(step);
4  │   for i=1:μ do
5  │   │   Produce ⌊λ/μ⌋ offisprings based on the i−th parrent;
6  │   end
7  │   Merge current P and new newP into one single population and choose
       the best μ solutions to replace the current population P.
8  end
9  best individual=best of the current population return best individual;
```

Algorithm 11: Basic ES($\mu + \lambda$)

Each new candidate solution y is created using Eq.2.6.

$$y_i = \begin{cases} x_i & \text{, with probability CR or if i=R} \\ a_i + F \cdot (b_i - c_i) & \text{, else} \end{cases} \tag{2.6}$$

where x, a, b, c are solutions of the current population, distinct from each other, y is the new solution, i is the $i-$th parameter and CR, F are the crossover weight and the weighting factor respectively. The former controls the amount of recombination, while the latter controls the amplification of differential variation. In detail, a high value of F leads to significant differential variation of the elements of solutions a, b and c, which increases the ability of the alorithm to escape local optima, while a low F leads to heavy dependence on solution a and therefore leads to quick local optima entrapment. On the other hand, an increased CR value leads to the construction of solutions that resemble much more x which can lead to diminished ability to combine good solutions to create better ones.

```
   Data: CR,F,POPSIZE
   Result: best individual
 1 Create initial candidate population P of size(POPSIZE);
 2 while termination criteria not met do
 3 │   newP = P+GaussianStep(step);
 4 │   for x ∈ Population do
 5 │   │   Choose three solutions a, b, c, distinct both from x and from each
 │   │       other;
 6 │   │   Pick up a randomly chosen variable Rfrom x;
 7 │   │   for xᵢ ∈ x do
 8 │   │   │   Pick a random number r ∈ [0, 1];
 9 │   │   │   if r < CR or xᵢ =R then
10 │   │   │   │   yᵢ = aᵢ + F(bᵢ - cᵢ)
11 │   │   │   end
12 │   │   │   else
13 │   │   │   │   yᵢ = xᵢ
14 │   │   │   end
15 │   │   end
16 │   │   ;
17 │   │   if y better than x then
18 │   │   │   y replaces x
19 │   │   end
20 │   end
21 end
22 best individual=best of the current population return best individual;
```

Algorithm 12: Basic DE

Table 2.4: Indicative Differential Evolution-based Image Registration Methods

Reference	Transformation	Description
[126]	Rigid	An early example of the use of DE in medical image registration, where, instead of $y_i = a_i + F(b_i - c_i)$, we have $y_i = a_i + F(b_i - c_i) + (best_i - a_i)$
[127]	Rigid	PSO is combined with DE and they are perfomed alternately. At odd generations, the new particles are formed using the original PSO, while at even generations, the DE is performed.
[128]	Rigid	DE is combined with Powell's Direction Set Method
[129]	Rigid	Parallel variant of DE. The mutation strategy here is DE/rand-to-best/1 and the candidate solutions are distributed across a grid. Initialization, evaluation and selection can be done in parallel, while the computational cost of reproduction is minimized in the mutation strategy by choosing randomly solutions from the neighbourhood of the one we want to mutate.

[130]	Rigid	DE is used to find a rough estimation of the solution and Regular Step Gradient Descent is used for refinement
[14]	Affine	DE is implemented as a distributed algorithm where several populations of candidate solutions are randomly created and evolving. Though a locally connected topology, each population is connected to a certain number of populations and exchanges canditate solutions with them.
[62]	Affine	Satellite Affine image registration using basic DE.
[131]	Affine	Multi-resolution variant, where DE is not using a single mutation method, but it autonomously adapts is mutation method according to its success to create better candidate solutions.
[132]	Rigid	DE is combined with artificial immune systems algorithm in order to speed up the registration process?
[133]	Rigid	DE is hybridized with Harmony Search
[134]	Rigid	Two new variants of mutation strategy in DE are presented: 1) a global mutation strategy where the same parameters F and CR are adapted dynamically and adopted by all the candidate solutions, 2) a local mutation strategy where the aforementioned parameters are associated to each individual candidate solution and are adjusted independently for each.
[135]	Affine	Orthogonal cross-over based on orthogonal experimental design [136], is compbined with DE. After each N iterations of DE, a randomly chosen percentage of the candidate solutions is used as parents for recombination and the best of the new offspring are inserted into the population. Experiments have proven that this variant converges faster to the global optimum.
[137]	Rigid	This variant of DE has two innovations: 1) F is not fixed but adaptive $F = F_0 \cdot 2^{\left(\exp^{\left(1 - \frac{G}{G+1-G_{max}}\right)}\right)}$, where G is the $g - th$ generation and G_{max} is the maximum number of generations. 2) Mutation strategy DE/best/2 is employed: $y_i = x_{best,i} + F \cdot (x_{r_1,i} - x_{r_2,i}) + F \cdot (x_{r_3,i} - x_{r_4,i})$. where r_1, r_2, r_3, r_4 are distinct from each other and from $best$.
[138]	Affine	The mutation strategy of DE is incorporated into the basic Artificial Bee Colony Algorithm in order to enhance the ability of the latter to escape local optima.

[139]	Rigid +Scaling	Mutation strategies DE/rand/1 and DE/best/1 are chosen alternatively in terms of the fitness values
[140]	Affine	Two innovations are incorporated in this DE variant: 1) Adaptive F ,2) DE/best/1 and DE/rand/1 mutation strategies are used by the good and the ba candidate solutions respectively.

2.5 Swarm Optimization Methods in Image Registration

Swarm Optimization methods, like evolutionary algorithms belong to the greater family of metaheuristic algorithms. However, unlike the later, which immitate aspects of Darwinian evolution, Swarm Optimization methods are inspired by aspects of 'collective intelligence' [141]. Collective intelligence is the result of the cooperation of large number of individuals, known as agents, in the environment. Its characteristics are decentralization, self-organization and distribution throughout the environment. Examples of collective intelligence are fish schools, flocks of birds and ant colonies.

2.5.1 Particle Swarm Optimization

Particle Swarm Optimization (PSO) [142] is an optimization method which solves a given problem via a population (swarm) of candidate solutions, known as particles. These particles move around in the search space according to simple mathematical formulae over their position and the velocity. The movement of each particle is affected by its local best position, but, also, it is guided towards the best known solutions in the search-space, which are updated, since other particles may find better potisions.

Starting from a initial population (of size $POPSIZE$) of random particles with initial position and velocity x_i, $v_i \in \mathbb{R}^N$, respectively and $i = 1, 2, \ldots, POPSIZE$ is the i-th particle. Let p_i be the best position of the i-th particle and b the best position found so far. Then, the basic PSO algorithm is:

$$v_i(j) = \omega v_i(j) + \phi_p r_p(p_i(j) - x_i(j)) + \phi_b r_b(b(j) - x_i(j)) \tag{2.7}$$

$$x_i(j) = x_i(j) + v_i(j) \tag{2.8}$$

where $\omega > 0$ is the inertial weight that balances the global and local search and ϕ_p, ϕ_b are positive constants. Numerous registration methods based on PSO have

been proposed some of which are presented in Table.

```
Data: ω,φ_p,φ_b,POPSIZE
Result: best individual
1  Create initial candidate population P of size(POPSIZE);
2  for i=1:POPSIZE do
3      Initialize velocity v_i U (−|b_up − b_lo|, |b_up − b_lo|);
4      p_i=x_i;
5  end
6  while termination criteria not met do
7      for i=1:POPSIZE do
8          for j=1:N do
9              r_p = U(0, 1), r_b = U(0, 1);
10             v_i(j) = ωv_i(j) + φ_p r_p(p_i(j) − x_i(j)) + φ_b r_b(b(j) − x_i(j));
11             x_i(j) = x_i(j) + v_i(j)
12         end
13         if f(x_i)>f(p_i) then
14             p_i = x_i;
15             if f(x_i)>f(b) then
16                 b = x_i
17             end
18         end
19     end
20 end
21 best individual=b return best individual;
```

Algorithm 13: Basic PSO

Table 2.5: Indicative PSO-based Image Registration Methods

Reference	Transformation	Description
[143]	Rigid	This version of PSO, apart from global best and local best, incorporates initial user guidance.
[127]	Rigid	DE with PSO
[144]	Rigid	Comparison of 3 PSO Variants
[145]	Rigid	A parallel variable neighborhood selection based PSO, with improved capability of escaping local optima.
[146]	Rigid+Nonrigid	The ω parameter is not constant but steadily decreasing in order to improve the performance of the algorithm: $\omega(t + 1) = \omega(t) + d\omega$, where $d\omega = \frac{d\omega_m in - d\omega_m in}{Max\_iter}$ and Max_iter is the maximum number of iterations.
[64]	Nonrigid	The particles are arranged in a dynamic hierarchy that is used to define a neighborhood structure. Depending on the quality of their so-far best-found solution, the particles move up or down the hierarchy. This gives good particles that move up in the hierarchy a larger influence on the swarm.

[147]		Compared to PSO, there is not the third term when particles update their velocities. In this case, information communication among the particles is weakening while the ability of local searching is reinforced. As a result, each particle has tendency to vigorously fly into its nearest extreme point so that the niche emerge. Since the calculation of the third term doesn't exist when particles update their velocities, the computation load of NPSO algorithm decreases a great deal. Moreover, its time complexity is only the linear function of swarm scale.
[148]	Affine	Preliminary step for nonrigid registration. PSO is combined with nonlinear least squares.
[149]	Rigid	Quantum-behaved PSO with disturbance implementation strategy in order to increase diversity in the population and thus avoid entrapment in local minima
[65]	Affine	Multi-resolution PSO
[150]	Affine	Hybrid PSO which incorporates two concepts of GA, crossover and subpopulation
[151]	Rigid	PSO with subpopulation and crossover
[152]		This variant combines Quantum PSO [153] with
[66]	Rigid	The velocity of each particle x_i is calculated as $v_i = w \cdot v_i + c_1 r_1(p_i - x_i) + c_2 r_2(b - x_i)$, where w is an inertia weight that adjustst the local and global optimization ability of PSO and it is not fixed but adaptive. When w value is bigger, it is easy to jump out of local search, and strengthen the global optimization ability. Conversely the smaller w value is beneficial to the local optimization ability enhancement and makes global optimization ability weaker.
[154]	Rigid	PSO with use of initial orientation for avoiding convergence of the swarm
[155]	Nonrigid	Quantum-behaved particle swarm optimization
[154]		Review
[156]	Rigid	PSO is used for locating the area of the global optimum and, then, Powell's method of minimization is used for fine tuning.
[157]	Rigid+Scaling	
[158]	Rigid+Scale	Use of fixed inertia weight for the adjustment of the particle velocity.
[159]	Rigid	inclusion of an unscented Kalman filter to guide particle motion, thus increasing the speed of convergence and reducing the likelihood of premature convergence

2.5.2 Other Swarm Optimization Methods in Image Registration

Although, PSO has been the most widely used Swarm optimization method for image registration, there are other swarm optimization methods that have been used for image registration such as:

- Firefly Optimization Algorithm [124, 160]

- Bacterial Foraging Optimization [161, 162]

2.6 Reduction of Computational Cost

Metaheuristic optimization methods, despite their success in image registration, have one disadvantage that is common to each and everyone of them: the need for repetive evaluations of the objective function for each candidate solution, that exists. This increases the computational cost, especially when the objective function is already computationally expensive. The most common means of dealing with it are the following ones:

- Parallelism

- Subsampling

2.6.1 Parallelism

Parallelism is widely used in solving computationally expensive problems. The idea is that several segments of the original serial code may be broken down into blocks of code that are independent from each other and, therefore, can be executed parallelly. Especially, because of the new technological advances in GPU programming, it is even easier to construct parallel image registration methods. The parallel methods are categorized with respect to:

- Device where parallel code is executed, i.e. CPU or GPU.

    - CPU [151]
    - GPU [117, 122, 163–165]
    - Both [5, 166–168]

- Standalone [5, 151] [163] or Grid [14, 28, 129, 167, 169–175]

- The degree of parallelization:

    - Optimization method [14, 163, 175]
    - Objective function [117, 122, 163, 166, 167]
    - Both

Table 2.6: Indicative list of Parallel methods of Image Registration

Reference	Type	Clusters	CPU	GPU	Optimzation Algorithm	Objective function
[169]	Rigid	Yes	No	No	Yes	No
[28]	Affine	Yes	No	No	Yes	No
[170]	Rigid	Yes	No	No	Yes	No

[176]	Affine	Yes	No	No	Yes	No
[172]	Non-rigid	Yes	No	No	No	Yes
[129]	Rigid	Yes	No	No	Yes	No
[177]	Non-rigid	Yes	No	No	Yes	No
[173]	rigid+Non-rigid	Yes	Yes	No	Yes	No
[174]	Affine	Yes	No	No	Yes	No
[163]	Non-rigid	No	Yes	Yes	No	Yes
[14]	Affine	Yes	No	No	Yes	No
[166]	Non-rigid	No	No	Yes	No	Yes
[175]	Rigid	No	Yes	No	Yes	No
[164]	Rigid	No	No	Yes	No	Yes
[151]	Rigid	No	Yes	No	Yes	No
[5]	Non-rigid	No	Yes	Yes	No	Yes
[117]	Rigid	No	Yes	Yes	Yes	Yes
[178]	Affine	No	Yes	No	Yes	No
[168]	Non-rigid	No	Yes	Yes	No	Yes
[165]	Non-rigid	No	Yes	Yes	Yes	Yes
[179]	Non-rigid	No	Yes	No	Yes	Yes
[122]	Rigid	No	No	Yes	No	Yes
[123]	Rigid	No	No	Yes	No	Yes
[180]	Non-rigid	No	Yes	Yes	Yes	Yes

2.6.2 Subsampling

In intensity-based image registration [2, 181],the pixel intensities of the images are compared via statistical measures such as Sum of Squared Differences [182–184], Cross-Correlation [181, 185–187] and Mutual Information [30, 32, 33, 37, 88, 188, 189]. Despite their success, the need to use all the pixels of the images increases the computational cost. This disadvantage is even more pronounced in 3D image registration. In order to deal with the increased computational cost, several methods have been proposed such as Parallelism, Algorithm Hybridization and Subsampling. In the case of subsampling, two main strategies [190] are employed:

- Downscaling of the image [191, 192] (e.g. by factor 1/2, 1/4, etc.). This can be very helpful in rigid and affine image registration, because the search space reduces.

- Use of a sample of the pixel intensities. Instead of using all the pixel intensities, a percentage of them is used (e.g. 50%, 25%). The pixels are selected either at random with uniform probability or sampled along a regular grid [33, 193].

The first way is widely used in rigid and affine image registration. The reason for this is the reduced search space in translation. For example, if the original image had dimensions $\mathbf{W} \times \mathbf{H}$, then the translation search space at X and Y axes is $(-\mathbf{W}, \mathbf{W})$ and $(-\mathbf{H}, \mathbf{H})$ respectively. However, if we start using a down-scaled version of the original image e.g. with factor 1/2, then the search space for the aforementioned parameters is reduced to $(-\mathbf{W}/2, \mathbf{W}/2)$ and $(-\mathbf{H}/2, \mathbf{H}/2)$ respectively. In short, if the downscale factor is $[k : k < 1]$, then the search space as a whole is reduced by k^2 and k^3 in 2D and 3D image registration respectively. In the second case [189, 194–197], instead of downscaling the image's size, we choose a sample of the pixel intensities in order to estimate the similarity. The subsampling factor (i.e. the percentage of pixels

we choose along the image dimensions), cannot be arbitrarily small without introducing local minima and deteriorating the image registration robust-ness [55, 198, 199]. The sampling is either deterministic or non-deterministic. In the first case, a single subset of pixels is selected and used throughout the registration process. The disadvantage of the deterministic subsampling is that convergence to the correct solution is not guaranteed due to the introduction of bias of the approximation error [111]. Apart from rigid image registration [38], subsampling has been used in non-rigid image registration, especially in medical imaging [111, 200].

2.7 Our Framework

As part of this doctoral dissertation, the work that has been done to improve automatic image registration is presented. In fact, this work is divided into three parts:

- Optimization Method

- Surrogate models

- Similarity estimation

2.7.1 Optimization Method

Image registration can be interpreted as a optimization problem where the goal is to maximize (minimize) an objective function that describes the (dis)similarity between two images. Therefore, we focused on finding better ways to improve the methods for optimization. Initially, we started from genetic algorithms and, later we moved onto Harmony Search which has not been used in image registration and is simpler, more efficient than Genetic Algorithms.

2.7.2 Surrogate Models Based on Machine Learning

It has been shown that Intensity-based methods are more robust than Feature-based methods, especially in multi-modal registration, using minimal (if any) image preprocessing at the cost of the increased duration. In order to deal with the computational cost, Surrogate models [201, 202] have been introduced in image registration.
Unlike previous methods such as parallelism and sub-sampling, which have been extensively used, Surrogate Models have not been used in image registration as a means of reduction of computational cost, despite having been heavily used in solving difficult and computationally expensive engineering problems.

2.7.3 Similarity estimation

The success of Mutual information and its derivatives as a similarity measure, especially in multi-modal image registration has paved the road for further research in other intensity-based similarity methods. Specifically, since Mutual Information of two random Variables **X,Y** is the Kullback-Leibler divergence (KL-divergence)of their joint Distribution and the product of their marginal Distributions (

$$MI(X,Y) = \sum_{x,y} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} = KL(X \cap Y, X \times Y) \qquad (2.9)$$

and KL-divergence is a special case of other more general forms of statistical divergences, one can construct other intensity-based image similarity measures. Among the most well known Statistical divergences are the following ones:

- F-divergences

- Tsallis' divergence

- Renyi divergence

Renyi's divergence is the most extensively used and has been used not only in rigid [37, 203–205], but also in non-rigid image registration [206–208].

To the best of our knowledge, however, no-one has studied how the robustness registration may be influenced from the α parameter of Renyi' divergence. In this thesis, this is studied together with the effect of subsampling through a series of experiments conducted on medical image pairs in order to examine how the registration accuracy is affected by the combination of the following parameters:

- Sub-sampling factor (i.e. the size of the image sample we use).

- α parameter.

In this work, we focused on similarity metrics based on Renyi' divergece and tried to examine their robustness depending on the value of the Renyi's parameter and the percentage of pixels we use to estimate the similarity of the images.

Chapter 3

Methodology

3.1 Rationale behind the research

Image Registration methods consist of the following important elements:

- Deformation model: What type of transformation must be applied in order to align the images.

- Similarity estimation: How the image similarity is estimated, given the image transformation. Are we going to use Features or Pixel Intensities? Feature-based methods are computationally cheaper, while Intensity-based methods are more robust especially when it is difficult to find salient features.

- Optimization method: How to reach the global optimum transformation, i.e. the transformation that will align the images. The method must be able to find the global optimum in each case.

Each method is constructed on the basis of different approaches to the above elements and according to the scientific area where Image Registration is applied. That means that under certain conditions or at a different scientific area an Image Registration method may fail. Despite the extensive research in Image Registration during the last twenty years, it remains an unsolved problem, i.e. there is no global method to match images acquired by the same or different sensors, under any lighting conditions for any transformation with minimal error in the minimum time and minimum, if any, interaction between the user and the method. Because of this and its extensive use in many scientific areas, there is a need for continuous research for new enhanced image registration methods that are able to deal with its issues.

3.2 Purpose of the presented work

The purpose of the presented work is to find the means to solve the unresolved issues or, at least, minimize the degree they affect the image registration problem. The issues of accuracy, robustness and computational cost are highly affected by similarity measures, transformation and optimization methods. In order to have an efficient and robust automatic method, accuracy must be high as well as optimization method must be automatic. Last but not least, the issue of the local optima convergence is affected by the optimization method. In order to deal with the issue of the automation, intensity-based similarity measures were used, since they require minimum, if any, image preprocessing. Because of their success and robustness in image registration, we decided to search for ways of improving them or, at least, minimize the effect of

their drawbacks. As it is mentioned above, intensity-based methods are able to perform better than feature-based ones when the images are poor in features as well as in multimodal registration. However, they are computational much more expensive from feature-based methods to the point that in previous decades, their computational cost was a prohibiting factor and, even now, they are still computationally expensive. The existence of many local optima in these methods is a problem that can be solved through the use of a robust optimization method.

Geometric transformation is an important parameter, which has been studied extensively, especially in the context of elastic deformation [40], so in this dissertation no research on it is presented. In detail, the presented work has focused on the following:

- Study of the similarity metrics, with the intention of finding better intensity-based measures which will solve (at least, partially) the issues of accuracy, robustness and computational cost.

- Study of the optimization methods, in order to discover better optimization methods that can cover the issues of accuracy, robustness, automation and local optima convergence.

- Introduction of machine learning as a means of reducing the computational cost.

3.3 How the main problems were addressed

The greatest problem with intensity-based methods is the computational cost. In our study, we tried to focus on a method that that is able to use that is able to find the optimal transformation while keeping the computational cost as low as possible.

3.3.1 Similarity metrics

The success of intensity-based metrics such as Mutual Information has started an ongoing research on other statistic measures as (dis)similarity measures in image registration. Mutual Information is based on Kullback-Leibler Divergence, which, in turn, is a special case of Renyi Divergence, a family set of statistical divergences. Although there are many other sets, this one has been most of all used in image registration, both rigid and nonrigid. In this thesis, we did several experiments on rigid registration in order to find how the parameter of Renyi divergence affects image registration.

3.3.2 Study of optimization methods

Our first steps were in genetic-optimization-oriented intensity-based image registration. Initially, we studied the robustness of elitist genetic algorithms, a variant of genetic algorithms that surpasses the basic algorithm in terms of robustness and speed. In our study, we did research on the effect of the mutation rate and the number of elites that are kept on the robustness of the aforementioned algorithm. Despite the initial progress we made, we tried to expand to other metaheuristic algorithms that can address the optimization problem more sufficiently. At the end, we choose Harmony Search [209, 210] for three reasons:

- Simple implementation.

- Efficient exploitation of the candidate solutions.

- Unlike other metaheuristic methods (e.g. Genetic Algorithms and PSO), this algorithm has not been used in image registration.

3.3.3 Introduction of Machine Learning as a means of reducing the computation cost

One disadvantage shared by all the metaheuristic methods is the repetitive objective function evaluations for each candidate solution that is created. There are several methods to reduce the computational cost in image registration such as parallelism and subsampling. In our case we tried to introduce machine-learning regression methods in order to partially substitute the similarity function with a cheap approximation. Although this approach, known as Surrogate models, has been widely used in many optimization problems, it has not been used in image registration.

3.4 Achievements

The Achievements are the following ones:

- In our initial study of Genetic Algorithms in Rigid Image Registration [211] we tried to discover the effects of Mutation rate and the number of elites in elitist genetic algorithm optimization. The results of our study have shown that the increase of either the mutation rate and/or the number of the elites make the optimization method more robust and less probable to diverge from the global optimum.

- Improving a variant of Harmony Search [209, 210] in the context of Rigid Image registration [212], thus successfully overcoming certain disadvantages we faced. Additionally, we devised a new termination criterion based on the mean and the standard deviation of the similarity estimation of the candidate solutions of our population.

- Successful introduction of Machine Learning as a means for reduction of computational cost [213] in Rigid Image Registration, which lead to the reduction of the computational cost by 47%.

- Further improvement of [212] and the creation of a new faster algorithm called Rehearsal-Harmony Search, which is faster and conducts better exploitation than the previous algorithm. This variant has been successfully used in Microscopy Image Stitching [214, 215] and affine image registration.

- In image comparison, we did research on similarity metrics based on Renyi divergence. We did research on the effect of Renyi's parameter α and the subsampling factor on the accuracy of our problem. Our study has shown than an increase of the parameter α leads to smaller registration errors.

3.5 Analysis of the achievements

Meta-heuristic optimization has been used in image registration, regardless of the transformation (rigid or non-rigid), comparison method (feature-based or intensity-based) and application areas (medical imaging or robot vision). In the following section, we present our initial approach in image registration using genetic algorithms.

Then, we proceed with the use of Harmony Search as a Meta-heuristic optimizer, which has never been thoroughly used in image registration. In this section, we present a brief description of the algorithm, as well as the changes that contribute to the improvement of the algorithm. The new algorithm variants were used in both in rigid and affine registration. Last but not least, the introduction of machine learning as a surrogate model in image registration is presented.

3.6 Rigid Registration

3.6.1 Genetic Algorithms in Image Registration

3.6.2 Description of the basic genetic algorithm

The performance of GA depends on following factors:

- Parameter Values (i.e. Crossover rate, Mutation rate and Population Size)

- Selection method

- Cross-over method

- Mutation-method

Parameter Values

The parameter values can affect the optimization method in numerous ways. A small population can easily get trapped into local optima, while a large one tends to slowly converge to the global optimum. Small mutation rates lead to slow convergence to the global optimum, while large ones tend to destroy good solutions before they can be exploited, thus leading to random exhaustive search. Last but not least, a high cross-over rate leads to premature convergence.

Data: Population Size **P**, Tournament Size (**k**)
Result: best individual
1 choose k (the tournament size) individuals from the population at random;
2 the best of the k random is inserted into the mating pool;

Algorithm 14: Tournament Selection

Selection process

The selection method must favour the fittest ones without the total exclusion of the worst solutions. After all, even the worst solution may have elements that can be used for reaching the global optimum. If the selection method favours only the fittest, then in the case when the current best solutions are near a local optimum, the exclusion of the rest solutions can lead to the use of only the ones at the local optimum and, therefore, to entrapment. There are several selection methods for such as:

- Truncation selection [216]: In this method, the candidate solutions are ordered according to their fitness (i.e. objective value estimation) and some proportion $p = 1/2, 1/3$,etc. of the fittest ones are selected and reproduced $1/p$ times. It is the least sophisticated of all the selection method and not often used.

- Tournament selection [217] : This method runs several tournaments among few randomly selected individuals from the population. In each tournament, the best is chosen for reproduction. The size of the randomly selected population affects the probability of choosing the least fit ones. The higher the number, the smaller the chance is for the least fit to be selected for reproduction, because there is greater chance to have to compete with fitter solutions. The advantages of this method are its simple and efficient coding, its parallelization and the selection pressure can be easily adjusted.

- Fitness proportionate selection or Roulette Wheel selection [218]: Each candidate solution is assigned with a probability of selection that is proportionate to its fitness. If an individual i has fitness f_i, then its probability of selection is $p_i = \frac{f_i}{\sum_{k=1}^{P} f_k}$. The advantage of the algorithm is that it allows even the worst candidate to be selected for reproduction since its probability p_i is not zero. Therefore, if that candidate has some good elements, they might be used during reproduction.

- Stochastic universal sampling [219]: Altough Fitness Proportionate selection may allow less fit candidate solutions to be used for reproduction, in the case where a member of the population has disdproportionately larger fitness (objective function value) from the rest, its assigned probalility of selection is significantly higher.

- Reward-based selection

Cross-over and Mutation Methods

The most common cross-over methods (Table 3.1) are the following ones:

- Single-point crossover [95]: At a randomly chosent point, the solotions split and the end-parts are exchanged in order to produce two offspring.

- K-point crossover: A generalization of the previous method where the split is done in K points.

- Even-odd crossover [220]: Even–odd crossover considers two operators: For even crossover, the even alleles of the offspring are taken from one parent and the odd alleles from the other parent. For odd crossover, the opposite is done.

- Uniform crossover [95]: In this method, the elements of the offspring are randomly chosen from those of the parents.

- Blend crossover [95]: Blend crossover chooses a new value for the offspring. Let x_i, y_i, $x_i < y_i$ be the i-th element of parents x and y respectively. Then the i-th element of the offsping is uniformly chosen from the interval $[x_i - a(y_i - x_i), x_i + a(y_i - x_i)]$, where a is a parameter of the method.

In the case where the elements of our solutions are numbers, the most common mutation methods are the following ones:

- Swap [95]: Two elements swap their values.

- Flip or Uniform mutation [95, 102–104, 220–222]: The element x_i mutates by sampling a random values from Uniform distribution whose boundaries are the value boundaries of the element.

51

- Gaussian mutation: Each element x_i is changed by a random value sampled from a Gaussian distribution $\mathcal{N}(0, \sigma)$ (i.e. $x_i = x_i + \mathcal{N}(0, \sigma)$).

- Quantum mutation [80]:

- Mixed mutation strategies [223–225]: Instead of using only one mutation strategy, two or more are combined.

- [57, 65, 78, 80, 83, 88, 94, 98, 102–104, 133, 190, 196, 222, 226–231]

Table 3.1: List of Crossover Methods

Crossover Methods	Description	
	Parents	Offspring
Single-point Crossover	$(x_1, \ldots, x_i, \ldots x_n), (y_1, ,, \ldots, y_i, \ldots, y_n)$	$(x1, \ldots, xi, yi+1, \ldots, yn), (y1, ,,,, xi, \ldots xn)$
K-point Crossover		
Even-odd Crossover	$(x_1, x_2, x_3, x_4), (y_1, y_2, , y_3, y_4, \ldots)$	$(x1, y2, x3, y4, \ldots), (y1, x2, y3, x4,)$
Uniform Crossover	$(x_1, \ldots, xn), (y1, \ldots, yn)$	$x_i' = random(xi, yi)$
Blend Crossover	$(x1, ,,,, xi, \ldots xn), (y1, ,,,, yi, \ldots yn)$	

Table 3.2: List of Mutation Methods

Mutation Methods	New value
Swap mutation	$x_i x_j$
Flip mutation	$x_i = Uniform(lower_i, upper)$
Gaussian mutation	$x_i = x_i + N(0, s)$
Quantum mutation	$x_i \leftrightarrow y_i$

3.6.3 Elitist Genetic Algorithms

As we described in Parameter values, although mutation helps escaping the areas of local minima, a high rate may destroy the good solutions of the current population and they may be replaced by less fit solutions before we can exploit the former ones. In order to maintain that the best solution(s) will not be quickly discarded (unless we find better ones), a variant of the Genetic Algorithms has been designed called Elitist Genetic Algorithms.

In this variant, instead of total replacement of the current population by the new population, the best solution(s), known as elite(s) is(are) kept and inserted into the new population, replacing some solution(s) of the new population (usually the worst one(s)). This strategy ensures that, even if no better solution is found, at least the best found so far are not discarded and lost. This can increase the speed of convergence to the global optimum.

Because of its ability to converge faster to the global optimum, it has been widely used in image registration [77, 79, 85, 93, 95, 113, 232–234]. Yest, as far as we know, no one has studied how the combination of the number of the kept elites and the mutation rate affects the optimization process and, therefore, the result of the rigid image registration. More analytically, we wanted to answer the following questions:

- How does the number of the elites affect the outcome of the optimization and, therefore, the registration process? Intuitively, the elites' purpose is to provide information regarding the (possible) location of the global optimum. Therefore, a higher number of elites may provide us with more information than the one provided by only one elite.

- Is the elites' effect on the optimization affected by the mutation rate? If so, how does their combination affect the registration process?

In our experiments, we studied how the behaviour of the Elitist Genetic Algorithm is affected for various values of the number of the Elites and the mutation rate. The results are presented in Chapter 4.

3.6.4 Harmony Search

Harmony Search is a relatively new population-based metaheuristic algorithm [209] that imitates the music improvisation process through which the musicians search for the perfect harmony.
More analytically, each harmony consists of notes and jazz musicians are improvising, in order to create the perfect harmony. They create new harmonies accoring to the following strategy:

- Choose notes from known harmonies and use them unchanged.

- Choose notes from known harmonies and they are slightly changed.

- Play random notes.

If the new harmony is better than the worst harmony of the population, then the former replaces the latter. Harmony Search explores the search space and creates new harmonies in a similar maner. The algorithm is presente in Table 4.

Data: h,par,$step\_size$,POPSIZE,MAX_ITER
Result: best individual
1 Create initial candidate population P of size(**POPSIZE**);
2 **for** $i=1:MAX\_ITER$ **do**
3   Create $new\_harmony$;
4   **for** $j=1:NVAR$ **do**
5   **if** $U(0,1) < h$ **then**
6   Choose $k = U\_INT(1, MAX\_ITER)$;
7        $new\_harmony$(j)=$P[k][j]$;
8   **if** $U(0,1) < par$ **then**
9          $new\_harmony$(j)+=$step\_size \cdot U(0,1)$;
10   **end**
11   **end**
12   **else**
13        $new\_harmony$(j)=$U(min_j, max_j)$;
14   **end**
15   **end**
16   **if** *new_harmony better than the worst harmony of P* **then**
17   Replace the latter with the new_harmony;
18   **end**
19 **end**
20 best individual= the best P;
21 return best individual;

Algorithm 15: Basic Harmony Search

As we see, the algorithm starts from a randomly initialized population P. Then, at each iteration it creates an new solution $new\_harmony$ in the following way:

$$new\_harmony[j] = \begin{cases} P[k][j] & \text{with probability h(1-p)} \\ P[k][j] + step\_size \cdot U(0,1) & \text{with probability hp} \\ U(min_j, max_j) & \text{with probability (1-h)} \end{cases}$$
(3.1)

where $NVAR$ is the number of the variables of the solution, h are the probability to use element of a solution from the population and par is the probability to slightly change an element from a randomly chosen solution. $(1-h)$ is the probability to choose a random value from the search space using uniform distribution $U(min_j, max_j)$, where min_j, max_j are the minimum and maximum values of the $j-$th variable. The reasons for choosing this metaheuristic optimization algorithm are the following ones:

- Unlike Genetic Algorithms and PSO, it has rarely been used in image registration.

- In GA, each solution $s \in \mathbb{R}^{NVAR}$ is constructed using two parent-solutions. If the number the chosen solutions for crossover is $N < POPSIZE$, then the number of possible pairs is

$$\binom{N}{2} = \frac{N!}{2!(N-1)} = \frac{N(N-1)}{2}$$
(3.2)

  Even if we use Uniform Cross-over method, from each pair we can create 2^{NVAR} possible pairs. Therefore, the maximum number of pairs that can be produced from the exploitation of the current population is

$$\binom{N}{2} \cdot 2^{NVAR} = N(N-1) \cdot 2^{NVAR-1}$$
(3.3)

  In PSO, a particle moves to a direction dictated by:

  - its current position
  - its best position
  - the position of the optimum found so far.

  therefore, each new position is created using n In constrast with GA and PSO, in Harmony Search each solution of the current population can contribute to the construction of a new solution. Since the $j-$th $(j = 1, \dots, NVAR)$ variable of the new solution can have any of the any of the $POPSIZE$ possible values $P[1][j], P[2][j], \dots, P[POPSIZE][j]$, then the number of possible pairs is

$$\underbrace{POPSIZE \cdot POPSIZE \cdot \dots POPSIZE}_{NVAR} = POPSIZE^{NVAR}$$
(3.4)

  This strategy has two advantages:

  1. Choosing becomes simplified and less time is spent in deciding who should be used for constructing new solutions. Because of that, the computational cost in Harmony Search is less than that of GA.

  2. The solutions of the current population can be more efficiently exploited and create a greater variety of solutions.

- Unlike GA, Harmony Search has two mutation methods. The random $U(min_i, \; max_j)$ with probability $(1 - h)$ and a step-size random walk $step\_size \cdot U(0, 1)$ with probability hp. The first one ensures the avoidance of entrapment at an area of a local minimum, while the second one can be used in order to explore and, therefore, exploit the area of a given solution.

- Last but not least, the performance of GA can be

Initially, we tried started experimenting with a variant of Harmony Search, known as Global-best Harmony Search (GHS) [235]. This variant has the same step as the basic Harmony Search. The difference lies in the second step, where the step-sized adjusted note or a randomly chosen harmony is replaced with a random note of the best harmony.

$$new\_harmony[j] = best[k] \tag{3.5}$$

where $best$ is the best harmony found so far and k a randomly chosen variable. We experimented with this variant in rigid image registration using satellite images (Fig. 3.1, 3.2). The results in the first example were not the expected ones, while in the second example, the optimization method managed to approach the area of the global optimum, albeit after a disproportionally large number of iterations. Due to the quality of the results, we decided to search for a way to improve this GHA in order to deal with the disadvantages it has. A detailed analysis is provided in the fourth chapter.

3.6.5 Surrogate modelling

In metaheuristic optimization, fitness approximation [236, 237] is often used in order to deal with demanding computational problems. The reasons for using fitness approximation are the following ones:

- The objective function is computationally expensive. In the case of the intensity-based methods, even if a sample of the images is used for estimation of their similarity, it contains hundreds of pixel intensities (if not more) which is enormous compared to the few dozens of features that are used in feature-based methods. The computational cost increases as the transformation method becomes more complex (e.g. projective is more expensive than affine, which is more expensive than rigid) or if we move from 2D to 3D image registration.
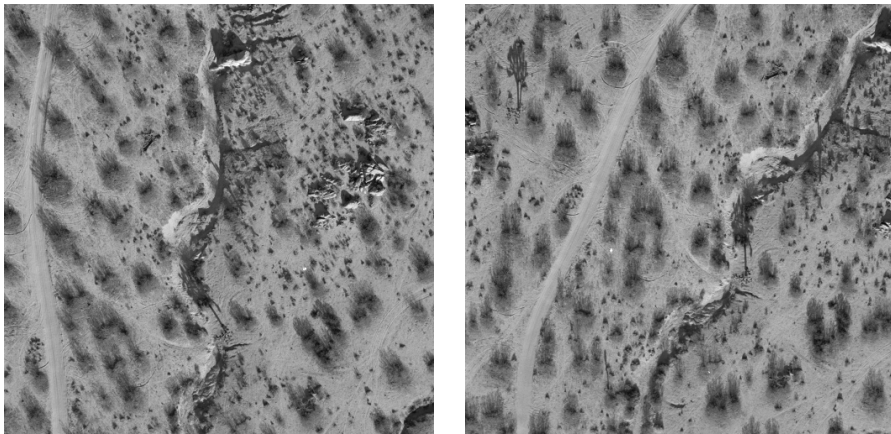
(a): b040   (b): b042

Figure 3.1: First Example of Satellite Images
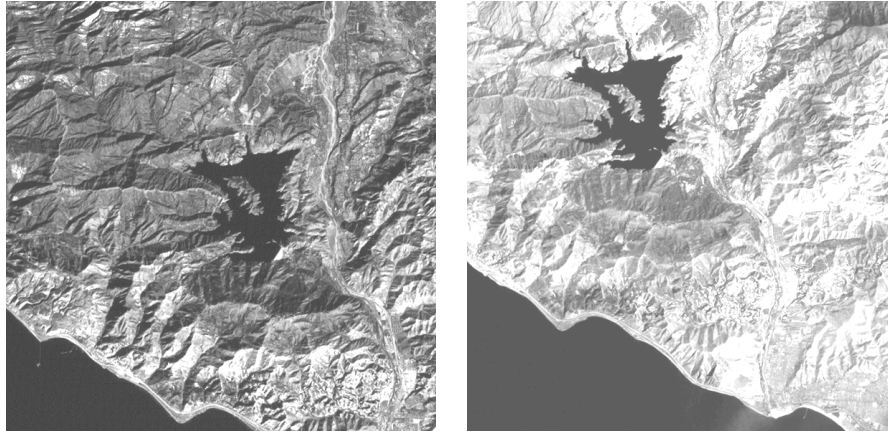
(a): casitas84   (b): casitas86

Figure 3.2: Second Example of Satellite Images

- The environment is of the algorithm is noisy. In the case of a non-noisy envir-onment, an individual solution can be approximated by averaging the function estimation of its neighbours. In a rather noisy environment such as that of an intensity-based method, this approach may not be useful.

- The landscape is multimodal. The idea is to construct a global model for ap-proximating and smoothing out the local optima of the original objective func-tion without changing the global optimum and its area. This can minimize the possibility of entrapment at the areas of local optima or, at least, enhance the escape from such areas.

Two are the traditional approaches in fitness approximation [238]:

- Problem approximation: This approach tries to replace the original statement of problem by an approximate one that is easier to solve. A good example is affine (6 and 12 parameters in 2D and 3D, respectively) registration of rigid image structures such as the human skull. It is known that affine registration includes the following transformations:

  - translation
  - rotation
  - scaling and
  - shearing

  However, in the example above, there is little, if any shearing effect. In this case, a good approximation of the solution can be found, if we try to reduce the problem and convert it from affine to similarity image registration (i.e. af-fine without shearing) which has 5 parameters and 9 parameters in 2D and 3D, respectively. Then, this approximation can be used as a starting point to find the global optimum in affine image registration. The reduction of the search space reduces the computational cost. An example of problem approxima-tion is [239], where complex perspective transformation is converted to simple scale transformation and translation transformation.

- Functional approximation: Here an alternate expression of the objective func-tion is constructed. A good(in terms of approximation) alternate and com-putational cheap expression replaces (at least, partially) the original objective function, thus reducing the duration of the optimization process.

In the case of intensity-based image registration, Functional Approximation could be the answer for reducing the computational cost. Also the smoothing of the local optima can reduce the number of iterations needed for escaping local optima. The most common methods used for functional approximation are:

- Fitness Inheritance [229].

- Polynomial models [236].

- Regression models.

- Machine Learning techniques.

In this thesis, we present the results of the experiments of functional approximation as a means of reducing the computational cost of image registration. In detail, the first experiments were conducted on medical images. Subsequently, new experiments were performed that compared different functional approximation methods using images of different sizes.

3.6.6 Stitching

As it is aforementioned, image stitching can be used in many different areas such as rapid detection of basal carcinoma cells [240] or depiction of biological structures [241]. In microscopy, there is great need for depicting all the details at the maximum resolution. The information that derives from the acquired super-resolution image is crucial for diagnosis and, therefore, treatment. An image stitching method for microscopy images that can be used by doctors and biologists must meet the following criteria:

- The method must produce high quality results with minimum error.

- It must be simple in use.

- It must be a generalized tool that can be used for every type of image.

Problems

Accurate image registration is the "sine qua non" factor in order to produce results of high quality. Here, we tried to stitch together a series of images acquired from a microscopic sample in a petri dish. Initially, we tried the grid stitching method [242] of FIJI [243], but we came up with the following problem: the petri dish (Fig.3.8) is circular. Suppose that we split the petri dish into overlapping horizontal series of overlapping images of equal size. The top and the bottom series have less images than the ones close to the center of the circle. Since the method of image stitching are using rectangular grids for image stitching, we came to the conclusion that this method is not suitable for such a problem.
A simple solution would be to do the following:

- First, stitch together the images in each horizontal series. After all, each horizontal series is a $1 \times N_i$ image grid, where N_i is the number of images in the i-th horizontal series ($i = 1, \cdots, M$).

- Second, stitch the results of the first step together. The results of the previous step are M images that are positioned on a $M \times 1$ image grid.
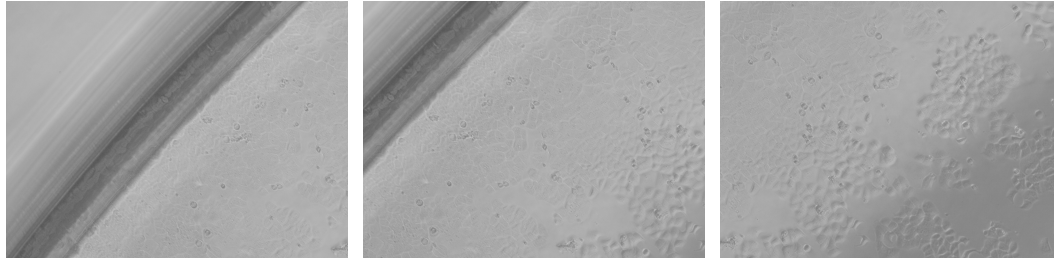
(a): Slice 1   (b): Slice 2   (c): Slice 3

Figure 3.3: Example of Satellite Images
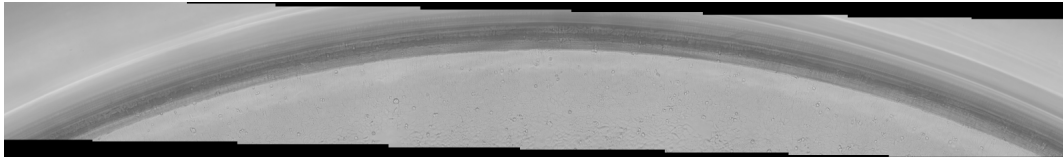
(a): Results with Tile Overlap equal to 0.2
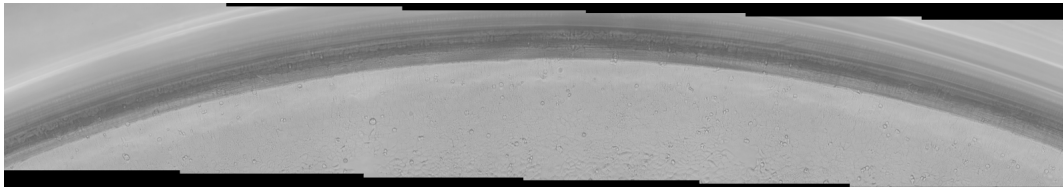
(b): Results with Tile Overlap equal to 0.3

Figure 3.4: Results of step one stitching of the first horizontal series
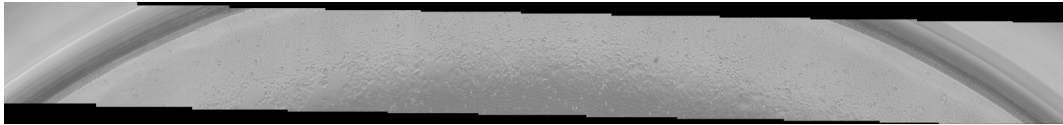
(a): Results with Tile Overlap equal to 0.2
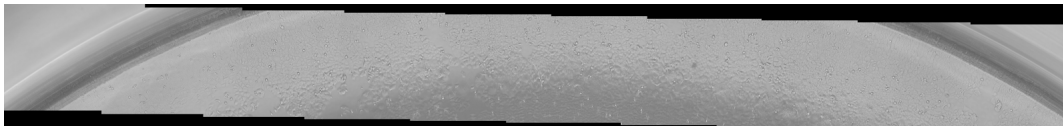
(b): Results with Tile Overlap equal to 0.3

Figure 3.5: Results of step one stitching of the second horizontal series

In both steps we used the grid image stitching method of FIJI (Fig. 3.7), Indicatively, we present some of the images we used in our experiments on brightfield images (Fig. 3.3) as well as some of the results of the experiments in Fig.5-7, using the above scheme .

As we see in the the results of Fig. 3.4,3.5, the increase of the Regression Threshold from 0.2 to 0.3 and from 0.3 to 0.75 respectively, instead of producing the same result, the first image on the right in both series was omitted. Although an increase of regression threshold can produce better results, it is obvious that it is not always easy to predict the results of a parameter's value variation, let alone when we deal with more than one. For a non- expert, this poses a even greater problem. This problem becomes more pronounced in the second step as it shown in Fig.3.6, where we change only the value of the regression threshold.

As we see in Fig.5-6, the problem of the choice of the correct parameter values is a very difficult, if not impossible, to solve. The parameters that align a series of images may differ significantly:
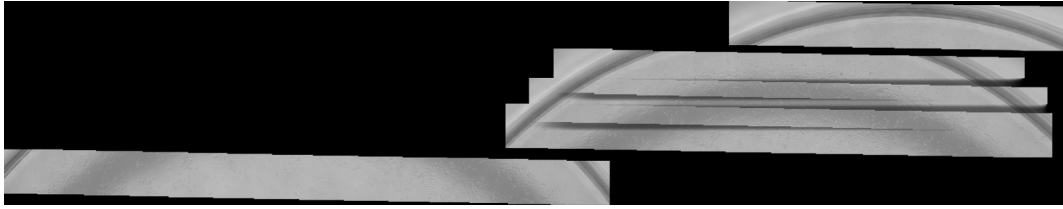
(a): Slive 1
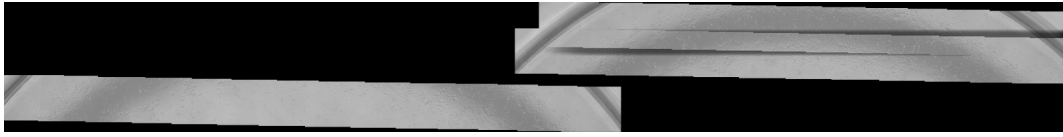
(b): Results of the second step of stitching the stitched series together with Regression Threshold equal to 0.25

Figure 3.6: Results of the second step of stitching the stitched series together with Regression Threshold equal to 0.25

- In the first step: Different horizontal series may require different parameter values.

- In the second step: Even in the case of the correct stitching of the horizontal image series, the new images vary in size which makes even more difficult the tuning of the variables, such as the overlap between them.

These disadvantages become more apparent in different image data, which make the whole process impractical, especially for non-experts. In the next section we present our innovations and how they solve the aforementioned problems.

New Approach

The ability of Mutual Information to align images using little, if any, preprocessing of the images and its robustness can be an asset. Initially, the new variant was intended to be used, but a closer study revealed the following disadvantage:

- If the new candidate solution is better than the worst member of our population, then the former replaces the latter.

- In the opposite scenario, the former is discarded.

In either case, a solution that could still be useful in the search for the global optimum is omitted. A second variant is constructed via the following assumptions inspired by the music sessions:

- When musicians are gathered to play together, they play in sessions.

- In each session, the musicians are creating new symphonies (solutions) using either:

  - Random notes (variables)
  - Symphonies from the previous session as well as those that are creating at the current one.

- At the end of the session, the best solutions are kept for the next session, while the rest are discarded.
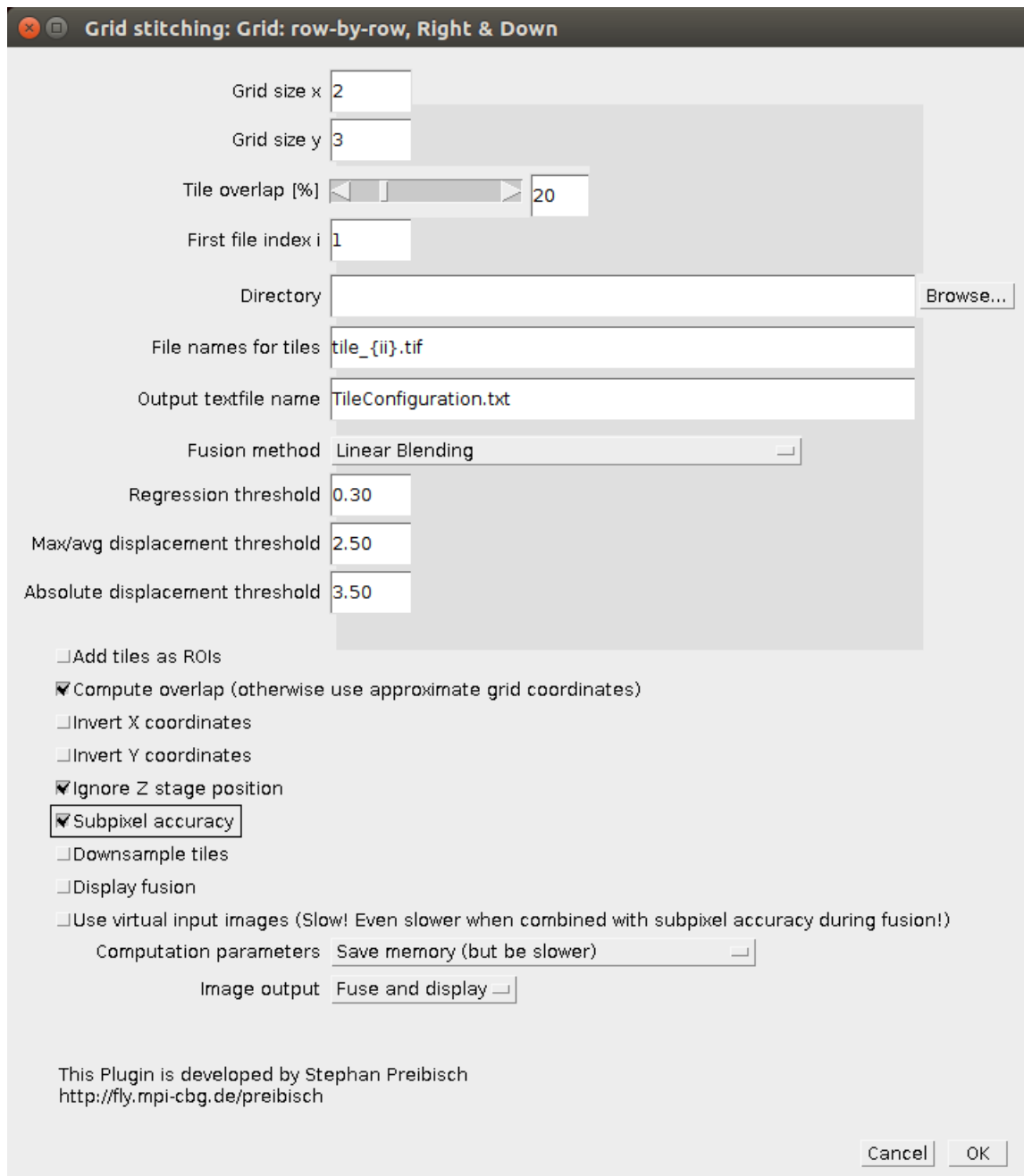
Figure 3.7: FIJI plugin
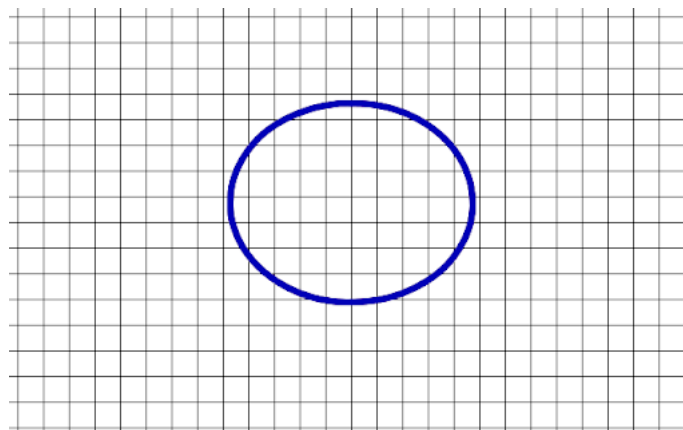
Figure 3.8: Circle on Grid

The second Harmony Search variant (called Session Harmony Search) is described in

Alg.16, where $POPSIZE$ is the population size, h_{min}, h_{max} are the mininum and maximum value of the h parameter, par_{min}, par_{max} are the mininum and maximum value of the par parameter. This variant differs from the first Harmony Search variant in the following aspects:

- Instead of creating a single candidate solution, in each iteration a new population of candidate solution is created and the from the current and the new population, the best $POPSIZE$ solutions are kept for the next iteration.

- In the original method and its variations, the new candidate solution is compared to the worst solution of the current population, replacing the last one if it is better. In the second variant, each new solution is integrated into the population, in order to be used in the construction of new solutions,thus expanding the availability of the elements that can be used for the construction of new solutions.

- After the construction of the a new candidate solution in basic Harmony Search, it is compared with the worst solution of the population. The computational cost of locating the worst solution is $O(POPSIZE)$. In the second variant, after the creation of the $POPSIZE$ new candidate solutions, they are grouped together with the current population and the $2 \cdot POPSIZE$ solutions are sorted according to their objective function estimation and the best $POPSIZE$ candidate solutions are kept. Using an effective sortin algorithm, the computational costs of sorting and replacement of the current population with the best $POPSIZE$ solutions are $O(2 \cdot POPSIZE \log(2 \cdot POPSIZE))$ and $O(POPSIZE)$, respectively. Therefore, the total computational cost of the second variant is:

$$O(2 \cdot POPSIZE \log(2 \cdot POPSIZE)) + O(POPSIZE) = \qquad (3.6)$$
$$= O(POPSIZE \log(2 \cdot POPSIZE)) + O(POPSIZE) = \qquad (3.7)$$
$$= O(POPSIZE \log(2 \cdot POPSIZE)) = \qquad (3.8)$$
$$= O(POPSIZE \log(\cdot POPSIZE)) \qquad (3.9)$$

3.6.7 Research on Image Comparisson

The success of Mutual information and its derivatives as a similarity measure, especially in multimodal image registration has paved the road for further research in other intensity-based similarity methods. Specifically, since Mutual Information of two random Variables **X,Y** is the Kullback-Leibler divergence (KL-divergence)of their joint Distribution and the product of their marginal Distributions (

$$MI(X,Y) = \sum_{i,j} \frac{p_{X,Y}(x_i, y_j)}{p_X(x_i)p_Y(y_j)} = KL(X \cap Y, X \times Y) \qquad (3.10)$$

and KL-divergence is a special case of other more general forms of statistical divergences, one can construct other intensity-based image similarity measures. Among the most well known Statistical divergences are the following ones:

- Tsallis Divergence

- F-Divergence

Data: h_{min},h_{max},par_{min},par_{max},$step\_size$,POPSIZE,MAX_ITER
Result: best individual
1 Create initial candidate population P of size(**POPSIZE**);
2 **for** *i=1:MAX_ITER* **do**
3   Create new population $NewP = P$;
4   **for** *i=1:POPSIZE* **do**
5          $h = h_{min} + (i/MAX\_ITER) \cdot (h_{max} - h_{min})$;
6          $par = par_{min} + (i/MAX\_ITER) \cdot (par_{max} - par_{min})$;
7   Create $new\_harmony$;
8   **for** *j=1:NVAR* **do**
9   Choose from $NewP$ 3 distinct candidate solutions $k, k1, k2$;
10   **if** $U(0,1) < h$ **then**
11                  $new\_harmony\text{(j)}+=P[k][j] + c(P[k][j] -$
                     $P[k1][j])\frac{(fun(P[k])-fun(P[k1]))}{abs(fun(P[k1])-fun(P[k2]))} + distribution(0, \delta)$;
12   **if** $U(0,1) < par$ **then**
13                      $k1 = second\_best, k2 = third\_best$;
                     $new\_harmony\text{(j)}+=best(j) + c(best(j) -$
                     $P[k1][j])\frac{(fun(best)-fun(P[k1]))}{abs(fun(P[k1])-fun(P[k2]))} + distribution(0, \delta)$;
14   **end**
15   **end**
16   **else**
17                  $new\_harmony\text{(j)}=U(min_j, max_j)$;
18   **end**
19   **end**
20   **end**
21   **if** *new_harmony better than the worst harmony of* P **then**
22   Replace the latter with the new_harmony;
23   **end**
24 **end**
25 best individual= the best P;
26 return best individual;

Algorithm 16: Second Harmony Search variant

- Renyi divergence

The reason for choosing Renyi divergence in our experiments is its wide use both in rigid and non rigid registration. As we mentioned above, there is no previous work studying the effect of Renyi's α parameter on registration accuracy. This becomes even more critical when considering the need for efficient subsampling to reduce computational cost. To address this issue, the experiments presented in this paper are designed to shed light on the following research questions:

- How does the parameter of Renyi's divergence affect the image registration error?

- How does the subsampling factor (i.e. the percentage of the pixel intensities we use) affect the image registration error?

- What is the relationship of the aforementioned factors regarding the image registration error?

To this end, a series of rigid image registration experiments are performed on image pairs
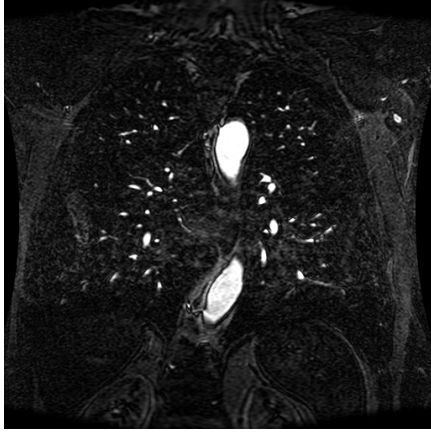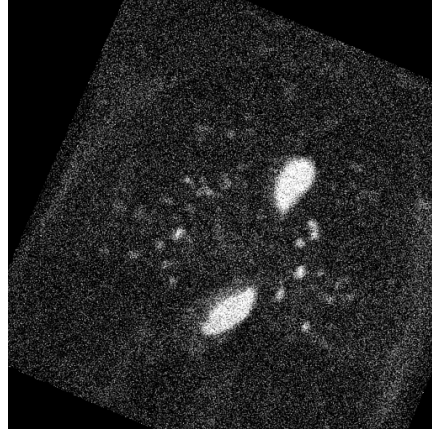
(a): Source Image   (b): Target Image

Figure 3.9: Example of image pairs
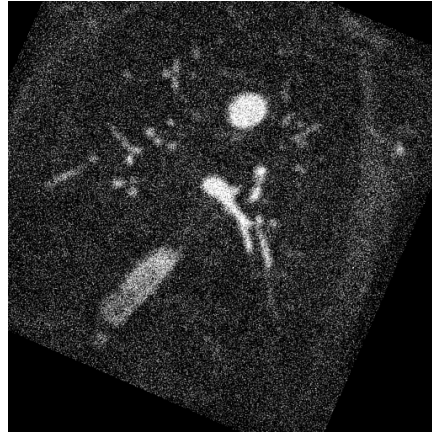
(a): Source Image   (b): Target Image

Figure 3.10: Example of image pairs

In the experiments conducted, twenty images were downloaded from a publicly available website [**?**]. Then, a copy of each one of the images is created and digitally processed in the following way:

1. Initially, the copies are blurred via a circular averaging filter.

2. Then, Gaussian noise is added to the blurred images.

3. Then, the images were transformed via rigid transformation.

For our experiments, the values of the parameters α and percentage (i.e. the Renyi parameter and the percentage of the sampled pixels) are the following ones respectively:

$$\alpha = \{0.1, 0.3, 0.5, 0.7, 0.9, 1.0, 1.1, 1.3\}$$

$$percentage = \{1, 2, 4, 6, 8, 10, 20\}$$

Non-deterministic sampling was used, in order to avoid any bias and assess how the parameters used (α, percentage), affect the registration error. Because of the non-deterministic sampling, it is expected to have different results in each run. Therefore, for each image pair, we conducted 100 experiments for any given parameter pair . Then for all image pairs, we align the images based on our metaheuristic algorithm [212] and then examine the rotation and translation for which the Renyi's divergence

(a): Source Image   (b): Target Image
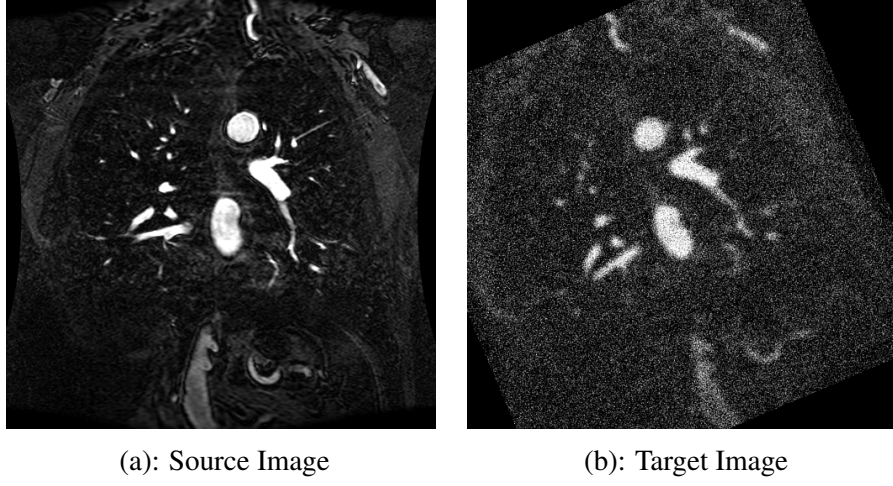
Figure 3.11: Example of image pairs

is maximized in order to calculate the absolute error. Renyi's divergence (eq.(**??**)), was implemented based on ITK tools [244]. In Tables I-III, we present the Mean Absolute Error (MAE) of the registration across all 20 image pairs with respect to rotation, translation along x and y respectively. MAE is defined in Eq. 7:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |err_i| \qquad (3.11)$$

where i is the i-th image pair, N is the total number of image pairs and err the registration error (radians for rotation and pixels for translations) with respect with the ground-truth transformations of the distorted images.

3.7 Affine

Affine registration has six parameters (9 in 3D) which makes the image registration more challenging than rigid. In Affine Registration, Elitist Genetic Algorithm, the first and the second variant of Harmony Search are compared. Also, in the case of the second variant, it was tested twice, both with and without the Support-Vector-Regression-based surrogate model, in order to study its impact on the results of image registration as well as the reduction of the computational cost. In our case, brain images were used, which were transformed via random affine transformation as we see in Fig.
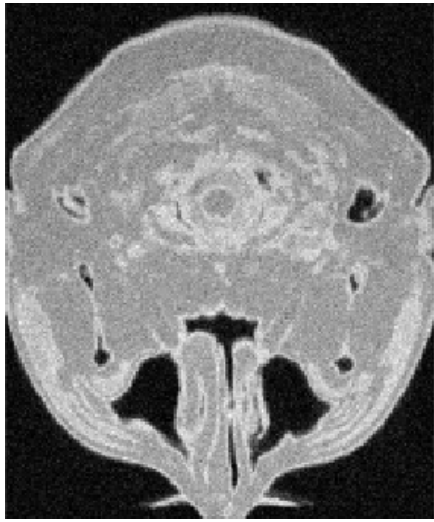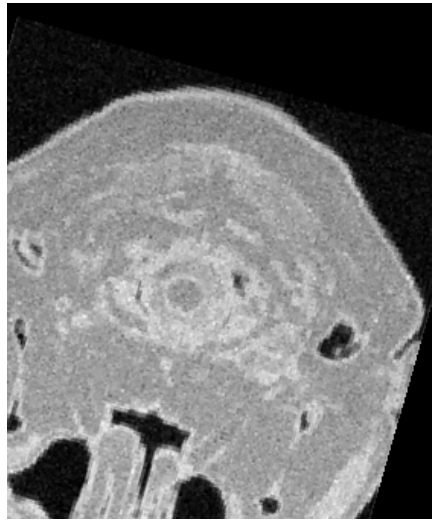
<div align="center">(a): Initial brain slice No. 5   (b): Initial brain slice No. 5</div>
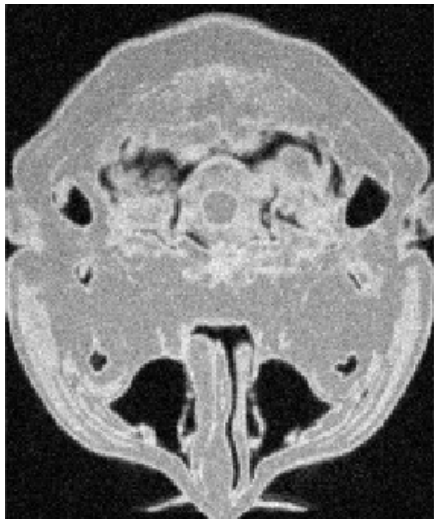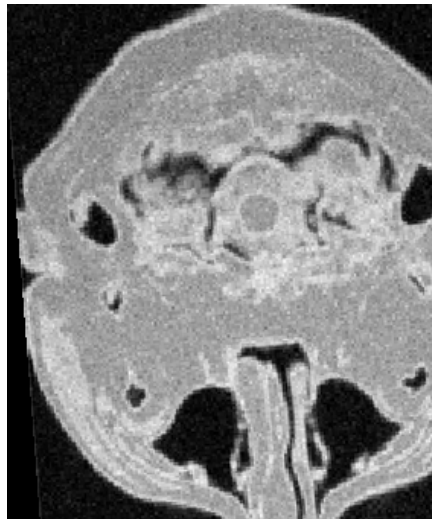
<div align="center">(a): Initial brain slice No. 10   (b): Initial brain slice No. 10</div>
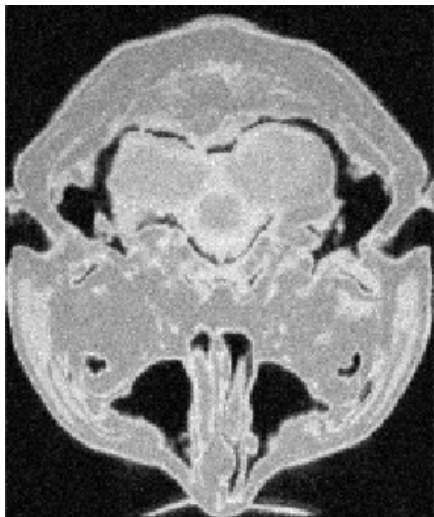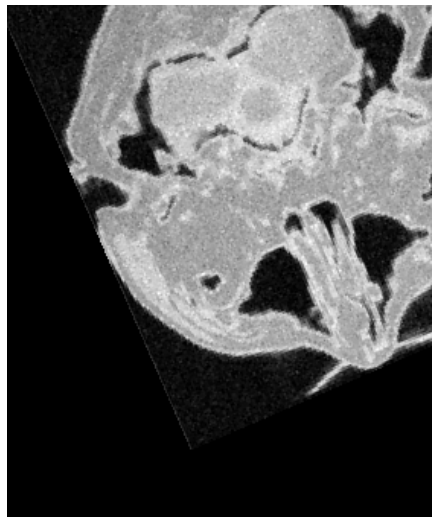
<div align="center">(a): Initial brain slice No. 15   (b): Initial brain slice No. 15</div>

Chapter 4

Results

Analysis, design, implementation and interpretation of results

The results of the study should be presented with as much detailed as possible. For quantitative and experimental studies, it must be clear how these results were obtained (i.e. what were the conditions which resulted in the findings). The discussion should provide interpretations behind the results and must explain how these fit into the existing body of knowledge.

In this chapter, we present the results of our research in detail as well as our findings regarding the new possibilities for image registration.

4.1 Rigid Registration

4.1.1 Elitist GA

As we mentioned in chapter 3, Elitist GA manages to converge quicker to the global optimum by replacing all but the best solution(s). Although this variant is a response to high mutation rates' disadvantage of destroying good solutions, no one has ever studied the effects of the number of the Elites and the mutation rate. In order to answer the questions, a series of experiments are executed in order to study the behavior of the IIR method with respect to the number of the elites that are kept and passed into the next generation and the mutation rate.

From [245] 25 image pairs were downloaded and used (representative pairs are shown in Fig.1) and in each one of them the target image is a contrast-changed version of the source image which is then randomly transformed (with the transformations shown in Table I). In each image pair, the parameters of the experiments we used were the following ones:

- Population Size: 100

- Maximum Number of Generations: 100

- Selection method: Fitness Proportionate Selection

- Crossover rate: 0.6

- Mutation rate: { 0.12, 0.14, 0.16, 0.18, 0.2 }

- Mutation method: For the first one third of the evolutionary process random uniform distribution is used. For the second one third cauchy distribution and is used and for the last one third we use gaussian distribution.

- Number of elites: 1, 2, 3, 4, 5

(a): Source I01   (b): Processed I01

(c): Source I02   (d): Processed I02

(e): Source I03   (f): Processed I03
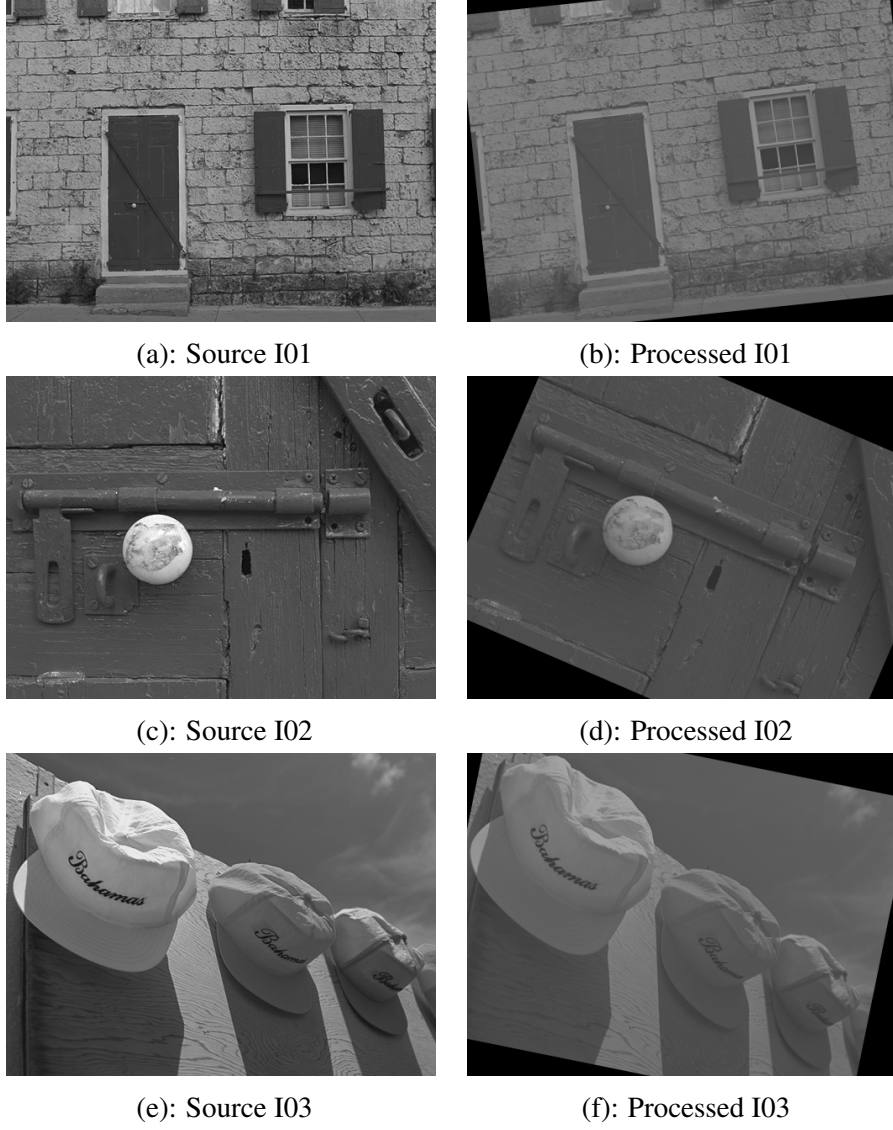
Figure 4.1: Example of image pairs

- Elitist method: Pass the elites unchanged

- Number of experiments per image pair per parameter pair : 100

In Figures 4.2-4.4., we present the results of the registration experiments using Elist GA optimization.

In Figures.4.5,4.6, we present the mean Mutual Information (left row) and its standard deviation (St. D. - right row) as a function of mutation rate (y-axis) and number of elites (x-axis) respectively for the first 10 image pairs. In detail, the average mean Mutual Information (left subfigure) and the average St.D of Mutual Information (right subfigure) for all pairs is presented. In Table 4.1, again for the first 10 image pairs (in the sake of space), we present the following:

- The number of the elites (# Elites) and mutation rate (Pm) for which the mean mutual information is maximized.

- The rigid transformation of the images, where rotation is measured in radians and translation is measured in pixels

- The mean absolute error of the optimum transformation found by the elitist algorithm using the best parameter values.

(a): Floating Image

(b): Reference Image

(c): Corrected Floating image

(d): Difference Image

Figure 4.2: First Example

(a): Floating Image

(b): Reference Image

(c): Corrected Floating image
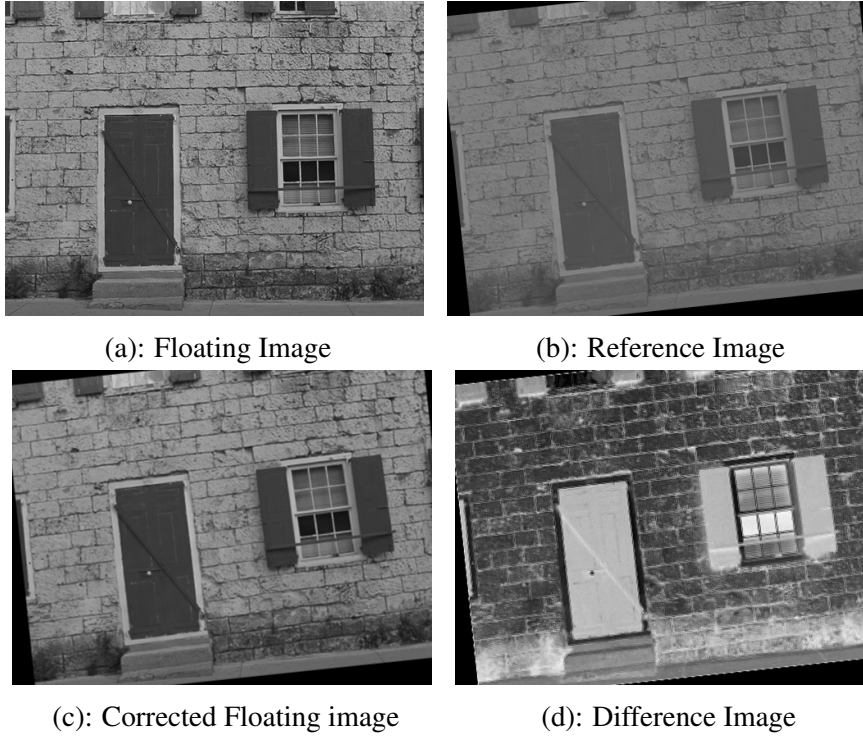
(d): Difference Image
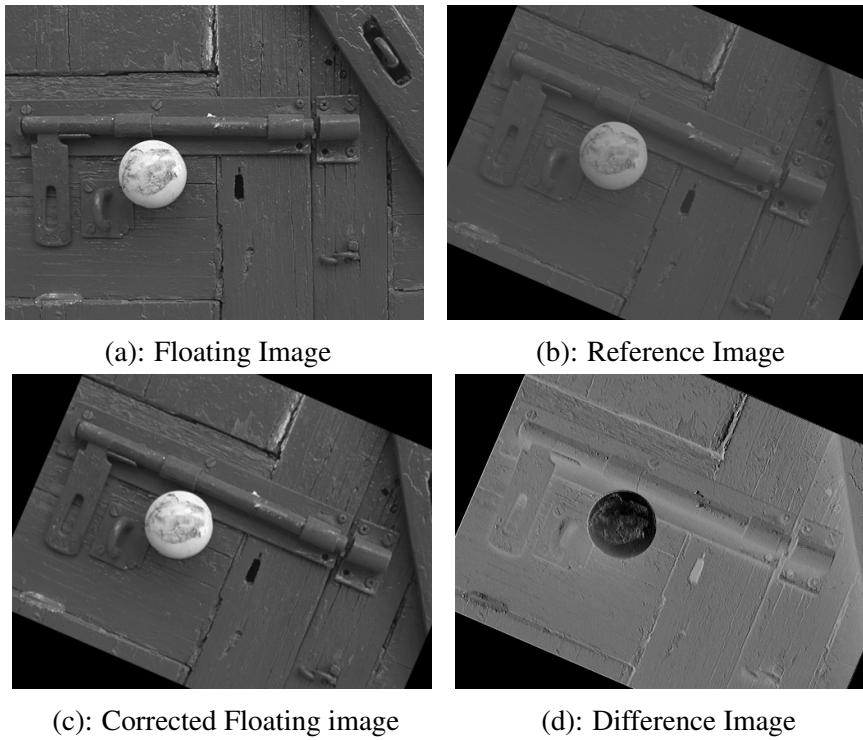
Figure 4.3: Second Example

The overall statistics of Pm and elites (and their pairs) which optimized MI for all image pairs are shown in the Table 4.2. It is obvious that in most cases high mutation rates and number of elites contributed to the maximization of Mutual Information across all image pairs (76% of image pairs achieved max MI with 5 Elites and 52% with 0.2 mutation rate). It is also interesting that 2 such combinations of # Elites and Mutation rates (0.2/5 and 0.18/5) resulted in maximizing MI in more than 60% of the images.

(a): Floating Image

(b): Reference Image

(c): Corrected Floating image

(d): Difference Image
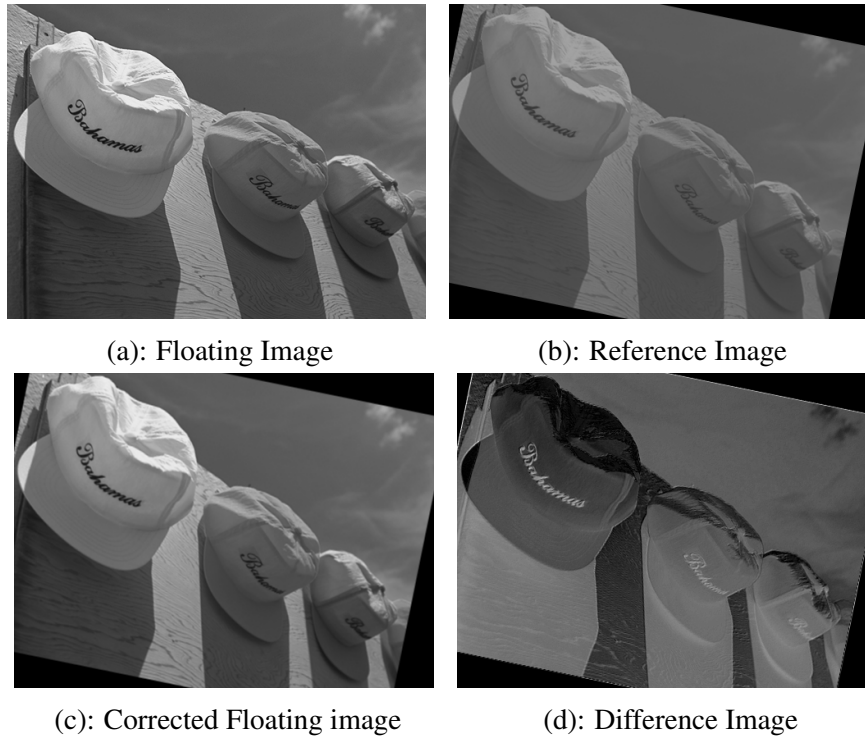
Figure 4.4: Third Example

Table 4.1: Statistics

No	Best Parameters		Transformation			Mean Absolute Error		
	P_m	# Elites	Rotation	X trans	Y trans	Rotation	X trans	Y trans
I01	0.2	5	-0.1047	8.4465	-8.3334	0.0000389	0.0119	0.0064
I02	0.2	5	0.4082	21.9414	8.6479	0.0000226	0.0058	0.0035
I03	0.2	5	0.2001	-15.4274	0.6511	0.0001648	0.0302	0.0141
I04	0.2	5	0.1767	-11.9486	44.4879	0.0001679	0.0048	0.0389
I05	0.18	5	-0.4715	7.3972	43.0496	0.0000236	0.0031	0.0088
I06	0.16	4	-0.4591	32.2138	-33.8409	0.0001381	0.0315	0.0229
I07	0.18	5	0.2192	28.6218	-6.9681	0.0000265	0.0123	0.0032
I08	0.2	4	0.4903	-23.1739	20.7755	0.0000631	0.0168	0.0098
I09	0.18	5	0.3669	-44.5967	12.8848	0.0002271	0.0047	0.0162
I10	0.18	5	-0.1571	-16.5133	48.0953	0.0000219	0.0075	0.0387

Table 4.2: Results

# Elites	5	4	3	2	1	
	76%	24%	0%	0%	0%	
Mutation Rate	0.2	0.18	0.16	0.14	0.12	
	52%	32%	16%	0%	0%	
Combined parameters	0.2-5	0.18-5	0.16-5	0.2-4	0.18-4	0.16-4
	36%	28%	12%	16%	4%	4%

4.1.2 First Approach in Harmony Search

Image Registration Using Global Harmony Search

As it is aforementioned in the previous chapter, we used Global-best Harmony Search (GHS), in order to see how well adapted it is for rigid registration. The parameters we used for our experiments were the following ones:
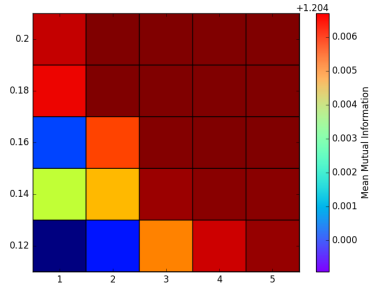
(a): Mean of MI of I01
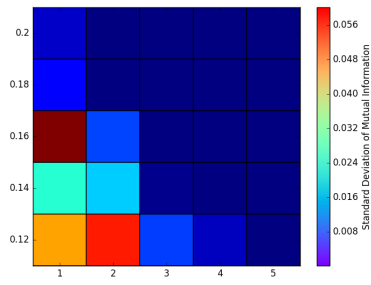
(b): St.D. of MI of I01
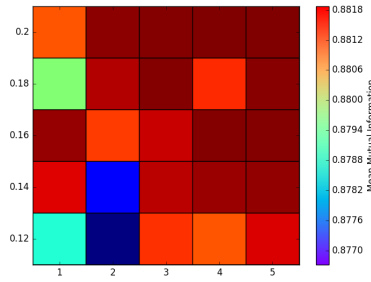
(c): Mean of MI of I02
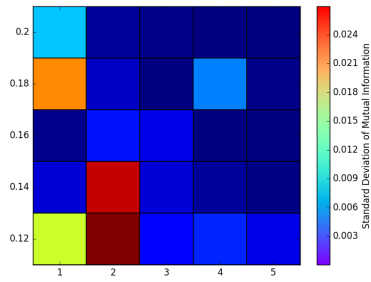
(d): St.D. of MI of I02
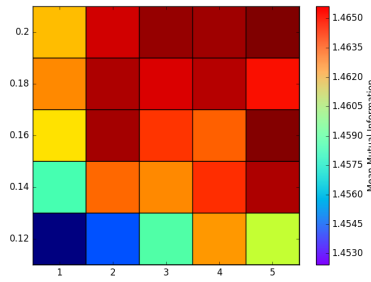
(e): Mean of MI of I03
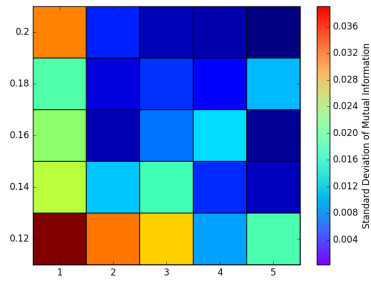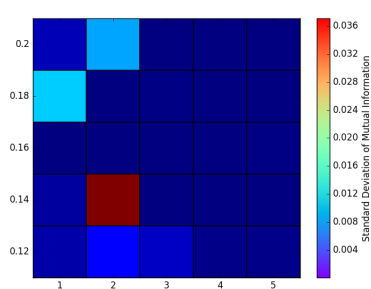
(f): St.D. of MI of I03
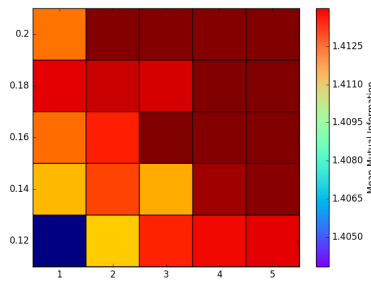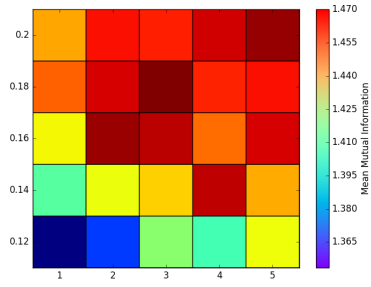
(g): Mean of MI of I04
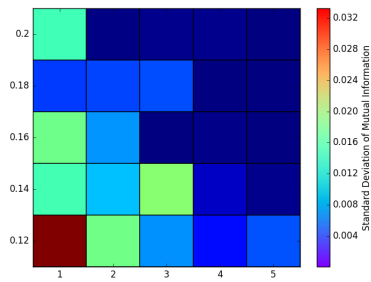
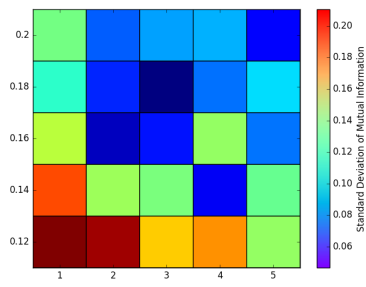(h): St.D. of MI of I04
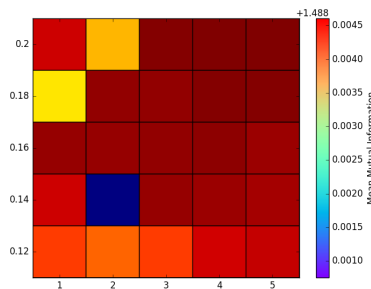
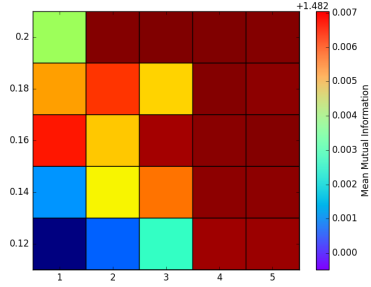(i): Mean of MI of I05

(j): St.D. of MI of I05
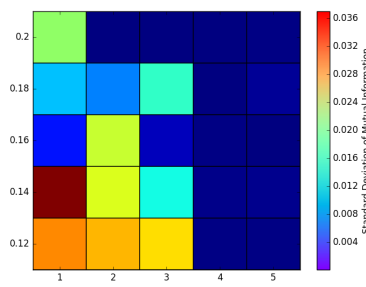
Figure 4.5: Images 1-5
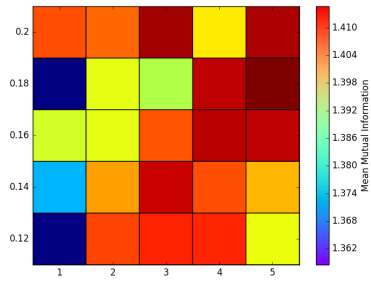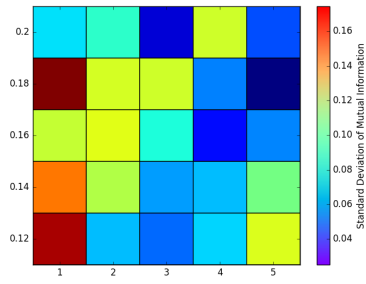
(a): Mean of MI of I06

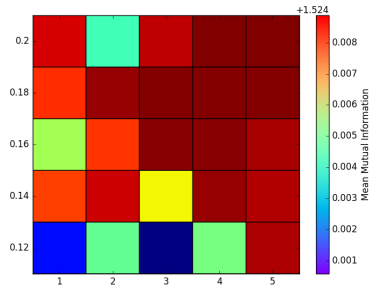(b): St.D. of MI of I06

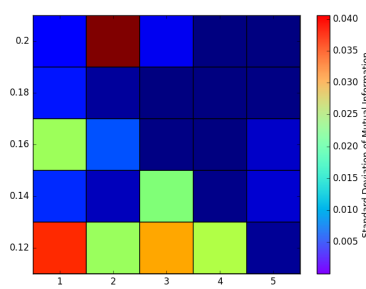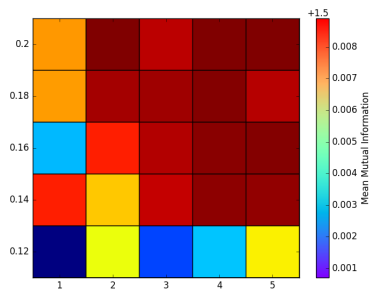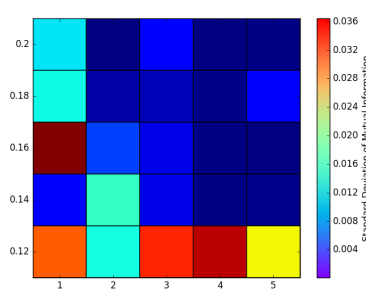(c): Mean of MI of I07

(d): St.D. of MI of I07

(e): Mean of MI of I08

(f): St.D. of MI of I08

(g): Mean of MI of I09

(h): St.D. of MI of I09

(i): Mean of MI of I10

(j): St.D. of MI of I10

Figure 4.6: Images 6-10

(a): Source image b040

(b): Target Image b042

(c): Transformed subfigure 4.11 (a)
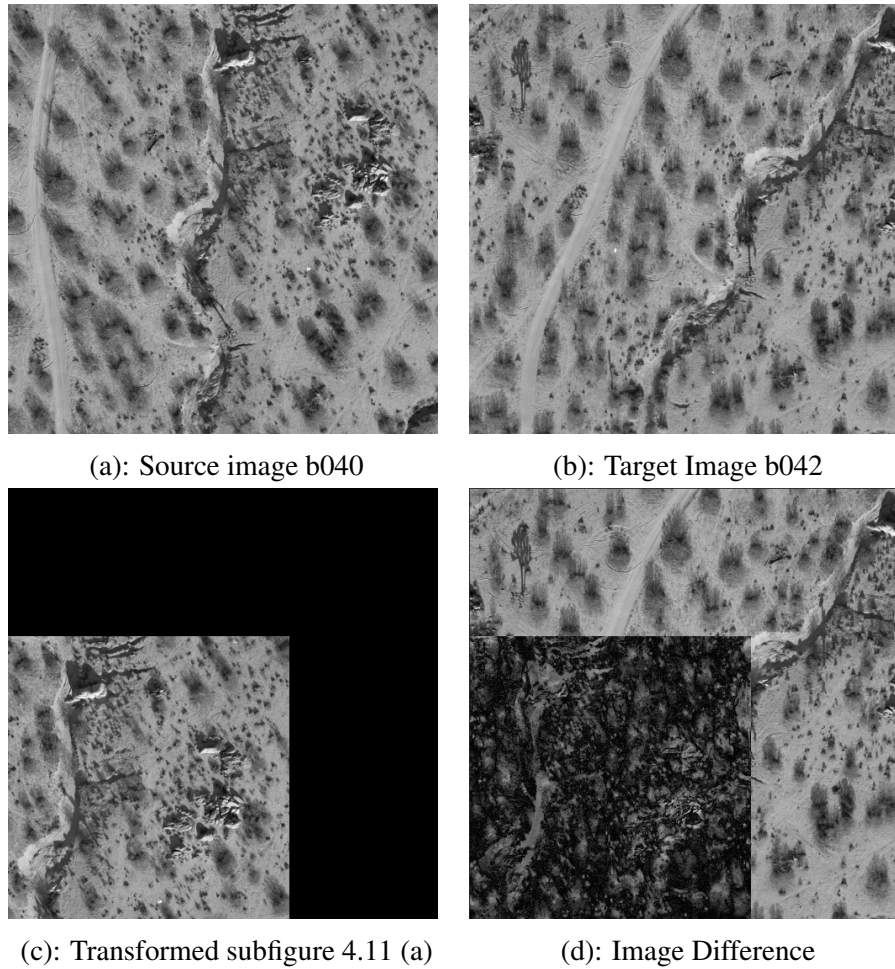

(d): Image Difference

Figure 4.7: First example

- hmrc: 0.7

- par: 0.3

- Population size: 20

- Number of Iterations: 240000

In fig.4.7,4.8, an example is presented that shows the results of the registration process. The registration was done using mutual information as a similarity measure. As we see in the Fig. 4.7, there is clearly a mismatch. On the other hand, in Fig. 4.8, the registration procedure seems successful. Nevertheless, another problem exists, which can be seen in Fig. 4.9. While in the first example b0 (green line) there is definetly an entrapment at an area of a local optimum, in the second example casitas (blue line) the registration method succeeds, albeit after a great number of iterations. The reasons of GHS being less effective than expected are the following ones:

- hmcr and par: These terms are fixed. An increased hmcr means that we choose to create solutions by using more often the harmony memory than exploring widely the search space. On the other hand, a small value of hmcr ensures avoiding entrapment at the areas of local optima, but it cannot exploit the memory in the case it has good solutions. A large value of par leads to frequent use of the best harmony, which may be a local optimum. If it is used to often, then the memory may soon be filled with solutions that are copies of it.

- The lack of perturbation: Using no perturbation in the algorithm reduces the ability of the algorithm to explore and exploit an area simultaneously. In the

(a): Source image casitas84

(b): Target Image b042

(c): Transformed subfigure 4.8 (a)
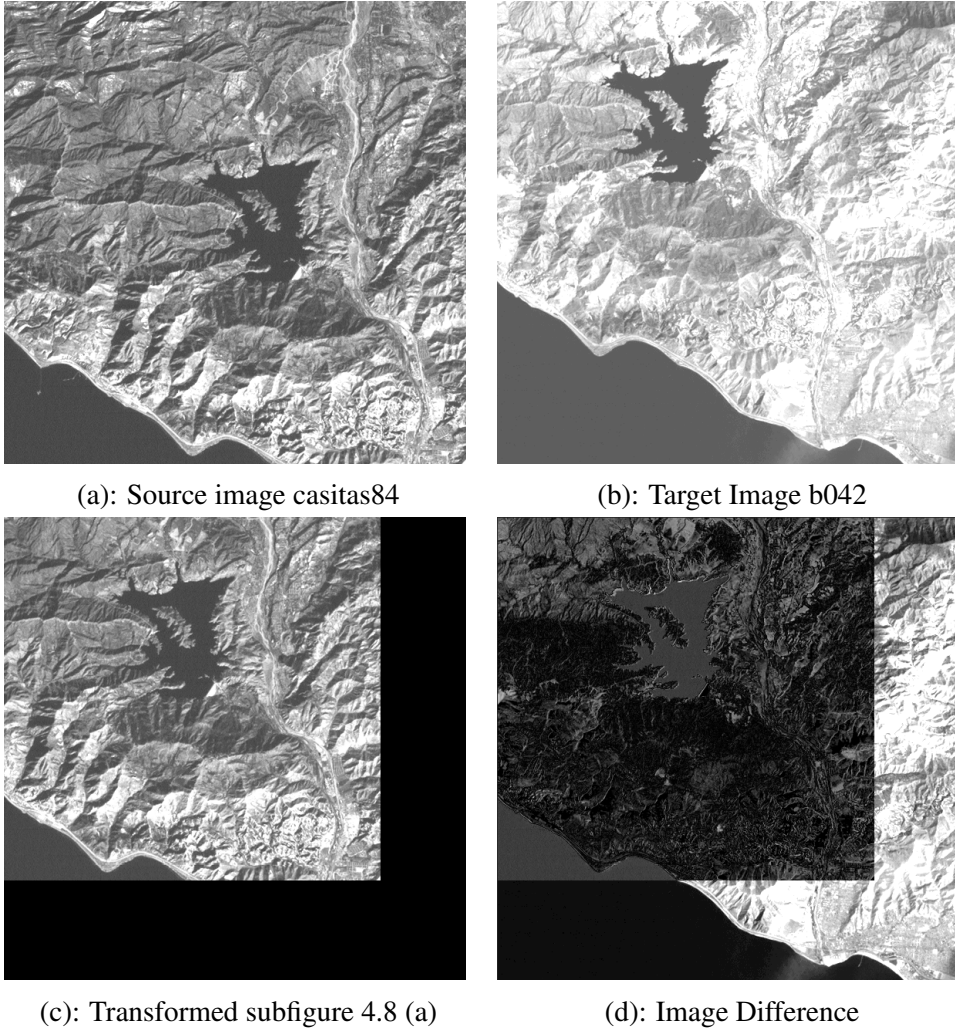

(d): Image Difference

Figure 4.8: Second example casitas

case of the second example, although the final result is successful, it takes too many iterations to escape the local optima and reach the global optimum.

- The algorithm does not exploit the information of the Harmony Memory efficiently. Although random mutations can lead to escape from local optima, it is could be better to guide the mutation towards better areas.

ALOPEX

ALOPEX (ALgorithn Of Pattern EXtranction) [246, 247] is an stochastic optimization that iteratively searches the values that optimize an objective function. This process combines the following:

- Bias feedback term, which uses the changes of both the variables and the function and tends to direct the optimization process to the area that has been proved successful in the recent past. In the case of the presence of the recent solutions at the area of the global optimum, the Bias Feedback term can converge faster to the global optimum.

- Stochastic term, which is a random number, generated for each variable in each iteration, and provides the opportunity to move each variable against the direction of recent success so that entrapment at the areas of local optima is avoided.
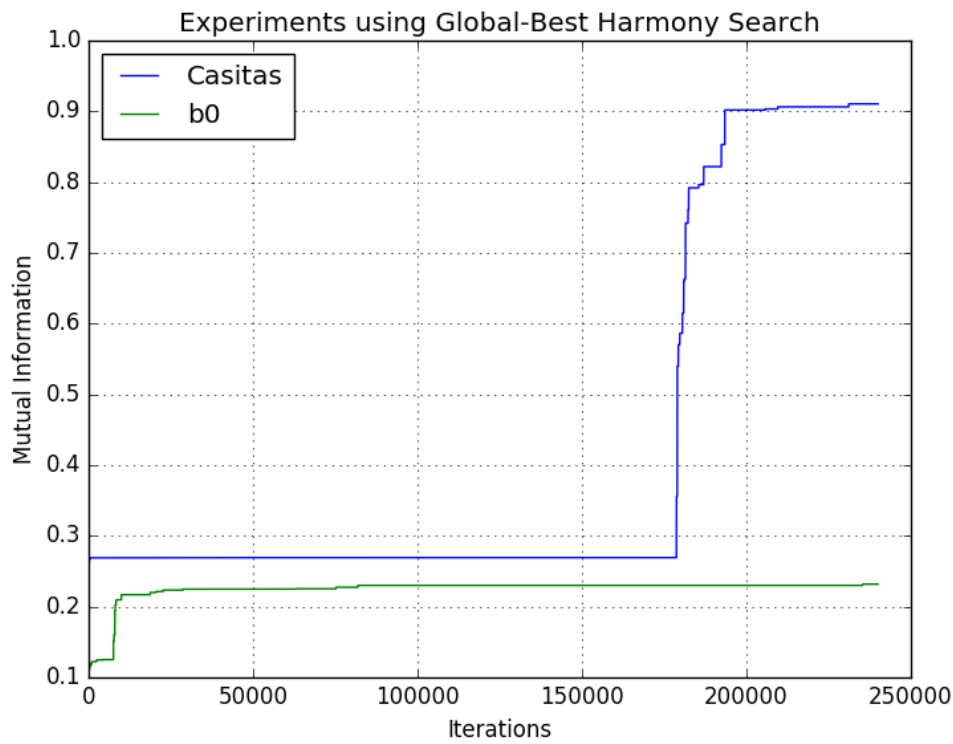
Figure 4.9: Mutual Information Graphs using GHS

A balanced use of both is critical for the success of the algorithm at converging to the global optimum. Too much reliance on the Bias feedback term reduces the algorithm to gradient descent algorithm, while its absence reduces it to random walking.
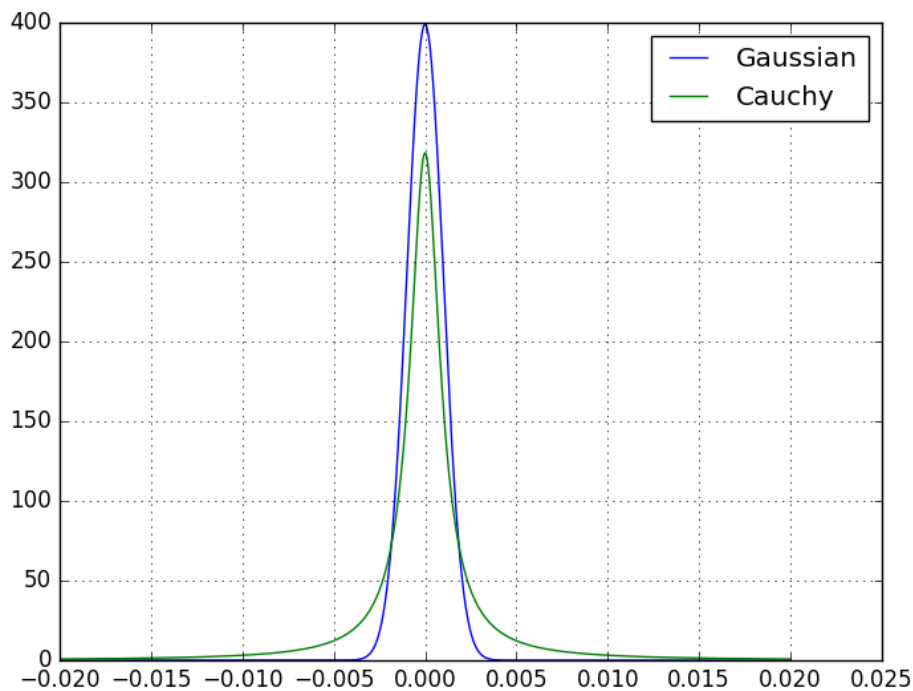

Figure 4.10: Cauchy and Gaussian Distributions with mean m=0 and standard deviations s=0.001.

Although the algorithm was initially used for correction of visual receptive fields [246, 247], its use has been expanded in several other fields such as image correction [248], cardiovascular of biomedical applications [249], pattern recognition [250] and neural networks [251]. Its main features [252] that are essential for the success of the algorithm are the following ones:

- Being stochastic and gradient-free, it depends on the variable values of previous steps rather than the functional form of the objective function

- The update variable is done simultaneously, which is not the case for other stochastic optimization methods such as Simulated Annealing. This process can be done in parallel, thus increasing it efficiency.1

- Unlike Simulated Annealing, the magnitude of the stochastic term does not depend on the amount it raises or lowers the response, which makes it ideal for moving between wide extrema. However, its magnitude that is significantly higher that that of the Bias feedback term, then the process can be reduced to random steps, much like the melting pot process in Simulated Annealing.

Novel Approach in Harmony Search for Rigid Image Registration

In order to deal with these problems a new variant is constructed (17):

- The $hmcr$ and par are not fixed: The parameter $hmcr$ starts at a lower value h_{min} because we want to focus on exploring the search space. Then, as the number of repeats increases, so does the parameter value because we want a gradual transition from exploration to exploitation until it reaches its highest value h_{max}. Similarly, the parameter par starts from a lower value par_{min}, because we do not want to depend too much on the current optimal solution because in the beginning this is most likely a local optimal. Finally, par steadily increases in order to exploit better the optimal solution until it reaches par_{max}.

- The optimal solution is exploited by integrating elements of the optimization algorithm ALOPEX IV [248].

- The most common termination criterion is the maximum number of iterations. To reduce computational costs due to the needless search for a better solution, we modified this termination criterion by examining in each iteration:

  - The absolute difference between respective evaluations of the best solution and the mean of the evaluation of the solutions in memory: If this difference is larger than a threshold/tolerance, then then new solutions can be created that are placed somewhere between the optimal solution and the rest. In fact, the closer they get to the former, they may find new solutions that are even better.

  - The standard deviation of the objective function evaluations: this shows how close the solutions are to the mean. If the standard deviation is higher than a threshold. it means that they can work together and find other better solutions.

- Unlike the original algorithm, which uses a uniform step in the secondary mutation, here Cauchy step is initially used and then it changes into Gaussian step. As we see in fig. 4.10, the bell-shape of the Cauchy distribution is wider than that of the Gaussian, which is an indicator of increased probability to further explore areas around the solution that are less probable for Gaussian

distribution. The latter is more useful once the area of the global optimum is found, where further exploration is most likely pointless and we need a more thorough exploitation via exploration at the area of the global optimum only.

The new algorithm is the following one is in 17.

Data: h_{min},h_{max},par_{min},par_{max},$step\_size$,POPSIZE,MAX_ITER
Result: best individual

1 Create initial candidate population P of size(**POPSIZE**);
2 **for** *i=1:MAX_ITER* **do**
3     $h = h_{min} + (i/MAX\_ITER) \cdot (h_{max} - h_{min})$;
4     $par = par_{min} + (i/MAX\_ITER) \cdot (par_{max} - par_{min})$;
5   Create $new\_harmony$;
6   **for** *j=1:NVAR* **do**
7   **if** $U(0,1) < h$ **then**
8   Choose $k = U\_INT(1, MAX\_ITER)$;
9             $new\_harmony$(j)=$P[k][j]$;
10   **if** $U(0,1) < par$ **then**
11   Choose $d = U\_INT(1, MAX\_ITER),\ d \neq k$;
12                 $new\_harmony$(j)+=$best(j) + c(best(j) - P[k][j])\frac{(fun(best)-fun(P[k]))}{abs(fun(P[k])-fun(P[d]))} + distribution(0,\delta)$;
13   **end**
14   **end**
15   **else**
16             $new\_harmony$(j)=$U(min_j, max_j)$;
17   **end**
18   **end**
19   **if** *new_harmony better than the worst harmony of* P **then**
20   Replace the latter with the new_harmony;
21   **end**
22 **end**
23 best individual= the best P;
24 return best individual;

Algorithm 17: First Harmony Search variant

The result of the use of the new variant are shown in Fig.4.11 as well as in the Graph at Fig.4.12. As we see in each of the experiments, the results were not only successful but also the convergence is much faster. Last but not least, the modified termination criterion caused the termination of the algorithm, preventing it from doing unnecessary computations and, therefore, reducing the computational cost.

4.1.3 Surrogate modelling

Initial Experiments on Medical Rigid Registration

As we explained in Chapter 3, a series of experiments were conducted for the reduction of the computational cost via functional approximation. Initially, Support Vector Regression (SVR) [253–256] was used for the construction of the surrogate model of Mutual Information in medical image registration. Although SVR has been used as an approximation method in image registration [257], it has not been used in the context of metaheuristic optimization in rigid image registration.
More analytically, we used the new harmony search variant in combination with SVR

(a): Source image b040

(b): Target Image b042

(c): Transformed subfigure 4.11 (a)
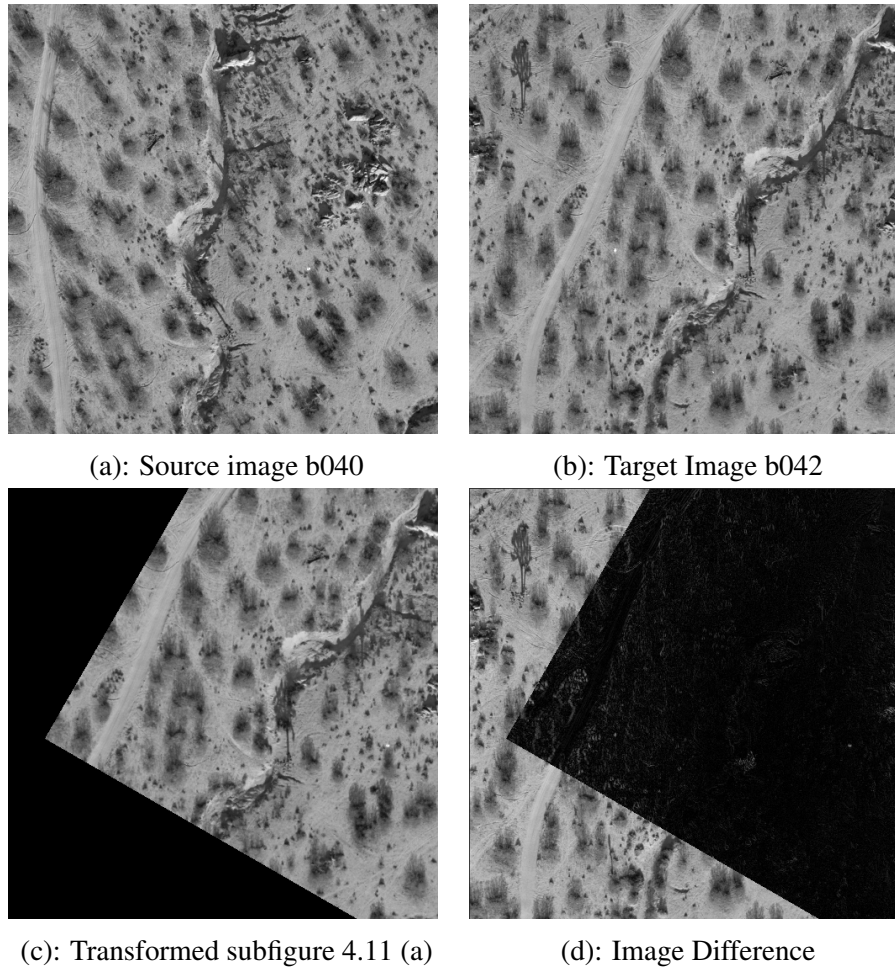
(d): Image Difference
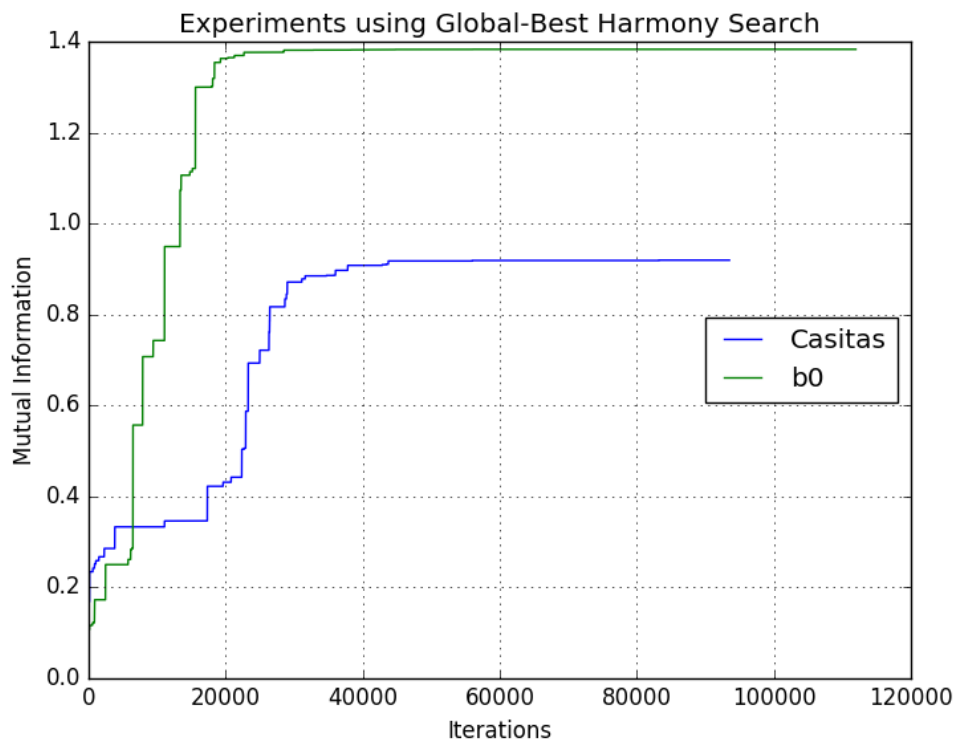
Figure 4.11: First example

Figure 4.12: Mutual Information Graphs using New HS Variant

as an approximation method and compared it with the original new harmony search variant. Due to the high non-linerity of Mutual Information, Gaussian Kernel ((4.1)) was used.

$$K(x, x') = \exp\left(-\frac{||x - x'||^2}{2\sigma^2}\right) \tag{4.1}$$

Initially, we create a random population of candidate solutions (whose size is P) and a training set (equal to the initial population) for the construction of the approximation. For every 2P new solutions, (P+1) of them are evaluated using the approximation method, while the rest (P-1) are evaluated using the original function and are inserted into the training set. When the training set reaches a certain size, the approximation is constructed anew in order to have an approximation method that is smooth and whose form (and therefore its global optimum), is (at worst) as close as possible to the real one. The approximation construction is repeated four times during the optimization process.

In the experiments, medical MRI images were used. These images (Fig. 4.13) are PD, T1 and T2 MRI images from several areas of the human body. In each case, the T1 and T2 is randomly transformed using rigid transformation. Then, a series of experiments are conducted both with and without the use of SVR. The experiments on each image pair (PD to T1 and PD to T2, from six different areas of the human body) are performed ten times using both the original method and the enhanced one in order to see the mean reduction of the image registration process as well as the results. Both methods succeeded in all the experiments for each image pair but the computational cost was consistently different as we see in indicative Figures.
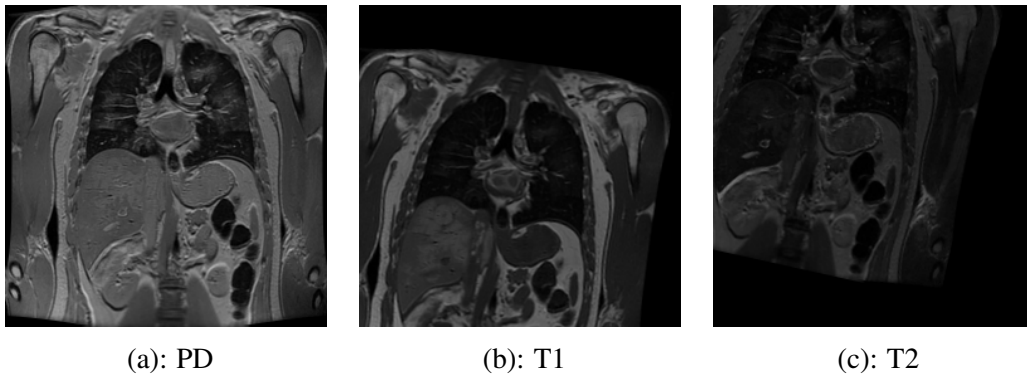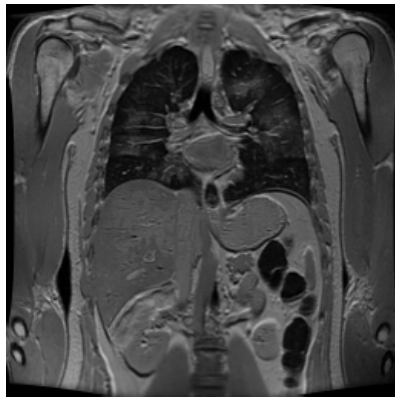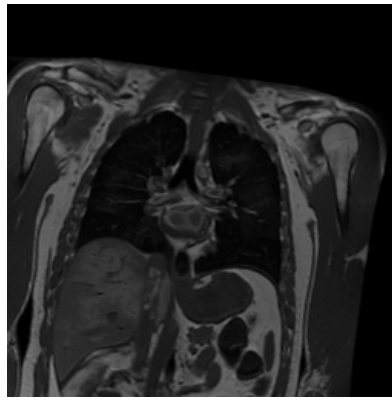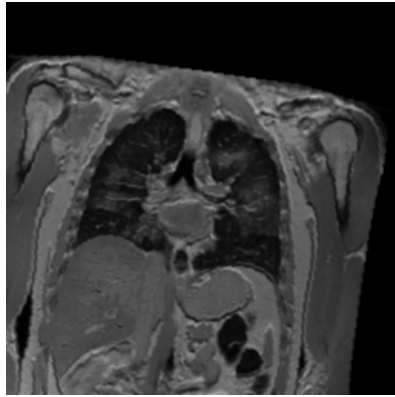

(a): PD   (b): T1   (c): T2

Figure 4.13: An example of images downloaded from `https://www.nlm.nih.gov/research/visible/mri.html`
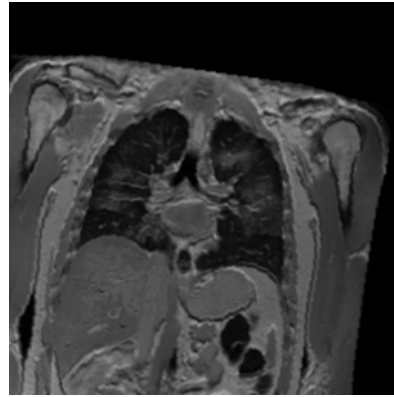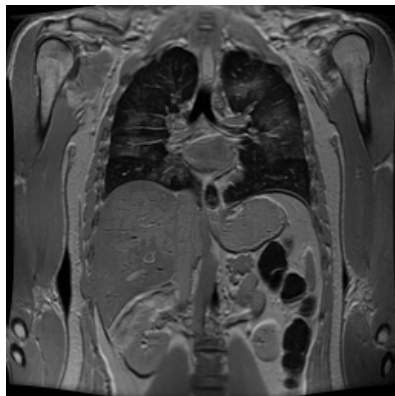
(a): Floating Image

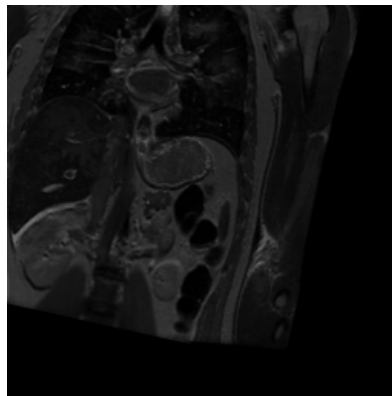(b): Reference Image

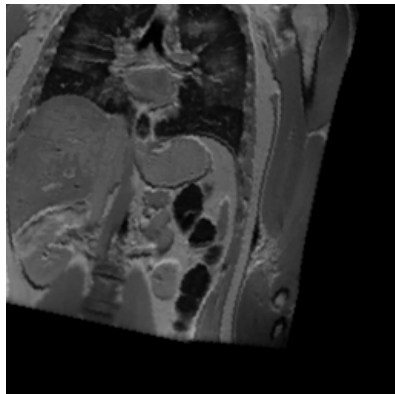(c): Result without SVR

(d): Result with SVR

Figure 4.14: Image Registration of MRPD and MRT1 at Abdomen area

(a): Floating Image

(b): Reference Image

(c): Result without SVR

(d): Result with SVR

Figure 4.15: Image Registration of MRPD and MRT2 at Abdomen area

(a): Floating Image   (b): Reference Image

(c): Result without SVR   (d): Result with SVR

Figure 4.16: Image Registration of MRPD and MRT1 at Pelvis area
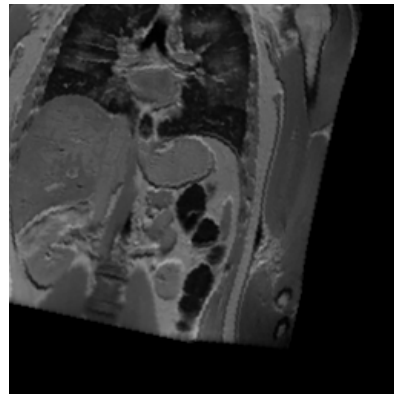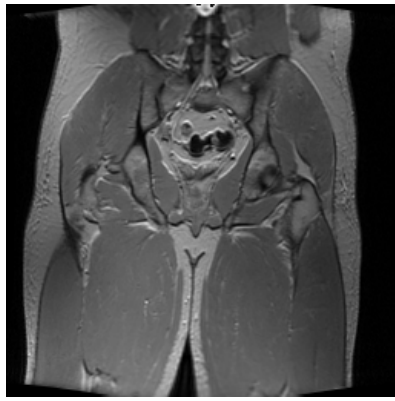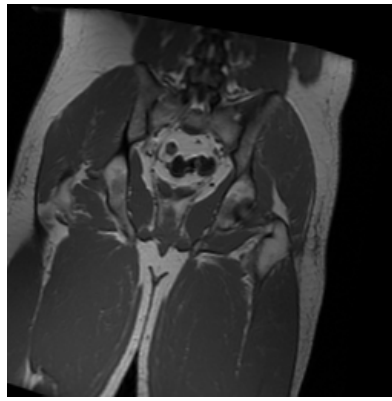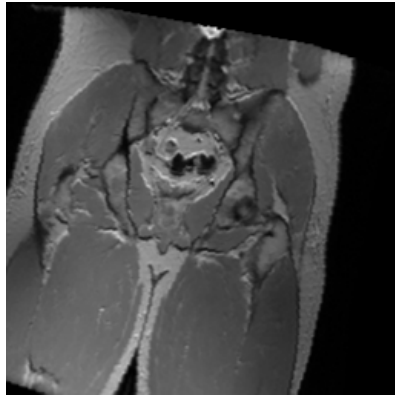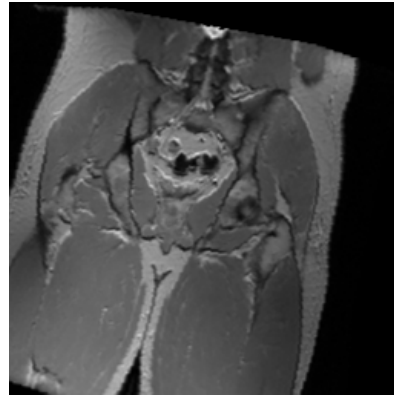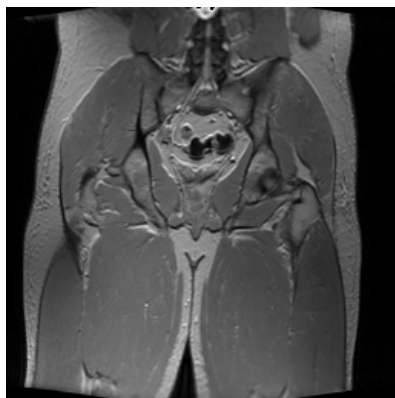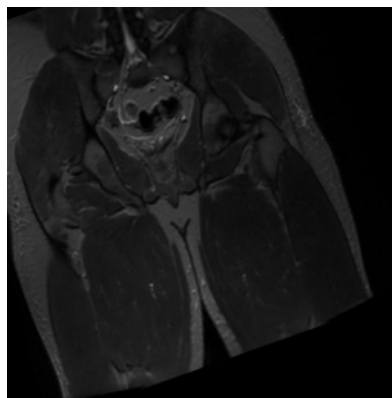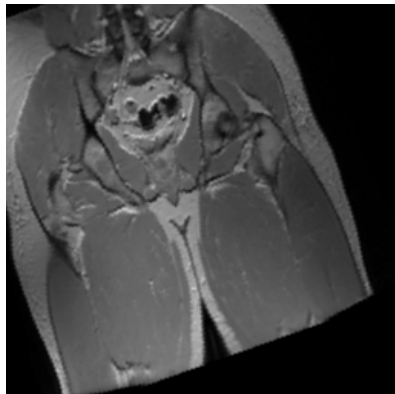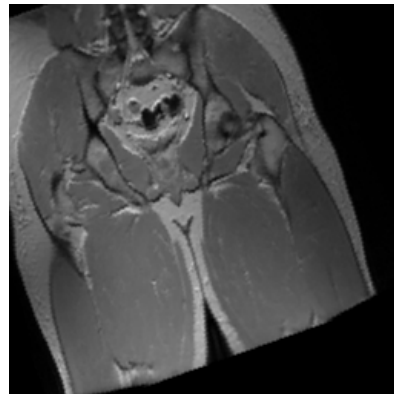

(a): Floating Image   (b): Reference Image

(c): Result without SVR   (d): Result with SVR

Figure 4.17: Image Registration of MRPD and MRT2 at Pelvis area

Table 4.3: Mean Duration of Image Registration in seconds

Image Pairs	Mean Duration	
	Original	Enhanced
Abdomen PD→ T1	2.884	1.871
Abdomen PD→ T2	2.909	1.745
Feet PD→ T1	3.129	1.989
Feet PD→ T2	3.048	1.907
Head PD→ T1	2.893	1.929
Head PD→ T2	2.947	2.066
Pelvis PD→ T1	2.926	2.042
Pelvis PD→ T2	3.146	1.918
Thigh PD→ T1	3.152	2.067
Thigh PD→ T2	3.067	1.960
Thorax PD→ T1	2.960	2.005
Thorax PD→ T2	2.957	1.991
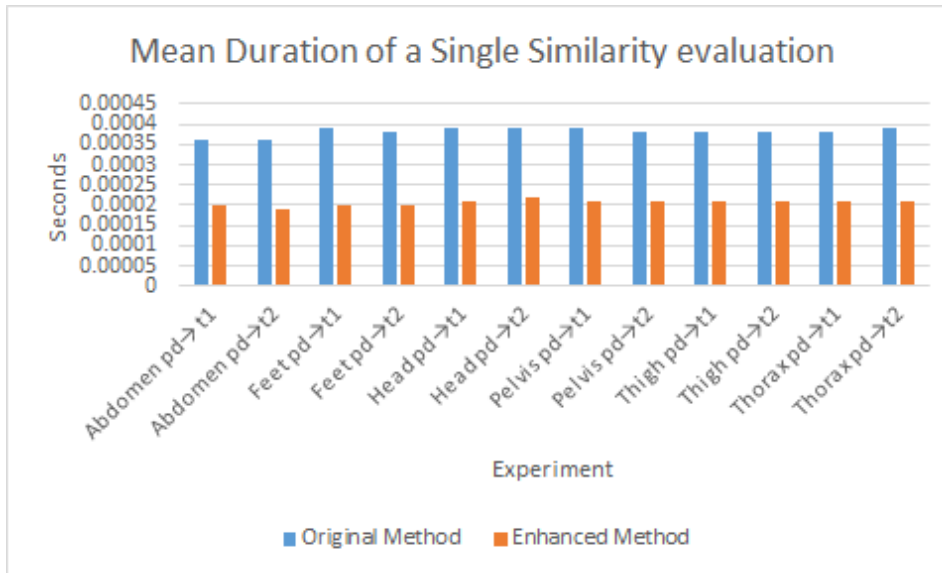
Figure 4.18: Mean duration of a single similarity evaluation

In Table 4.3, the mean duration of the method both with and without the SVR are presented. In Fig. 4.18 the mean duration in seconds of a single evaluation for both methods is presented. The reduction of the duration is up to 40.03%, while the mean duration of a single similarity evaluation (if we include the time needed for the construction of the approximations) is reduced to 45.55%.

Further experiments using Images of varying sizes

Previously, we used SVR as an approximation method for minimizing the computational cost of Mutual-Intentity-based Rigid Image Registration on Medical Images, whose size is 256. A series of experiments were conducted in order to estimate the which one of the following approximation methods can are best suited for Mutual Information where 5 different approximation methods are used:

- Kernel Recursive Least Squares algorithm (krls) [258]

- Kernel Ridge Regression (krr trainer) [259, 260]

- Multi-layer Perceptron (mlp)

- Radial Basis Function Network (rbf_network_trainer) [261]

- Epsilon-Insensitive Support Vector Regression (svr trainer) [262]

A series of medical [263] and non-medical images [245, 264] were used in the experiments. In these tests, a randomly generated sample of 1500 transformation was used for each image pair. The quality of the approximation methods is validated by executing the following steps:

(a): Source Image   (b): Target Image
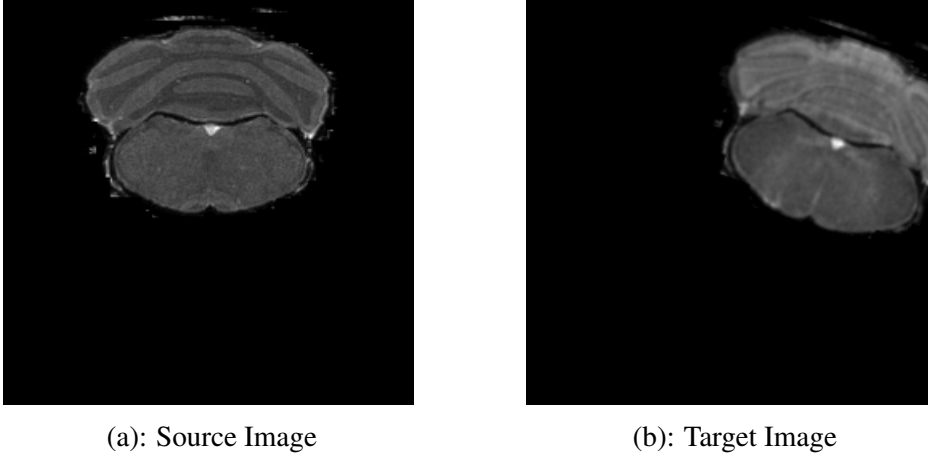
Figure 4.19: Mouse brain image slice 110

(a): Source Image   (b): Target Image
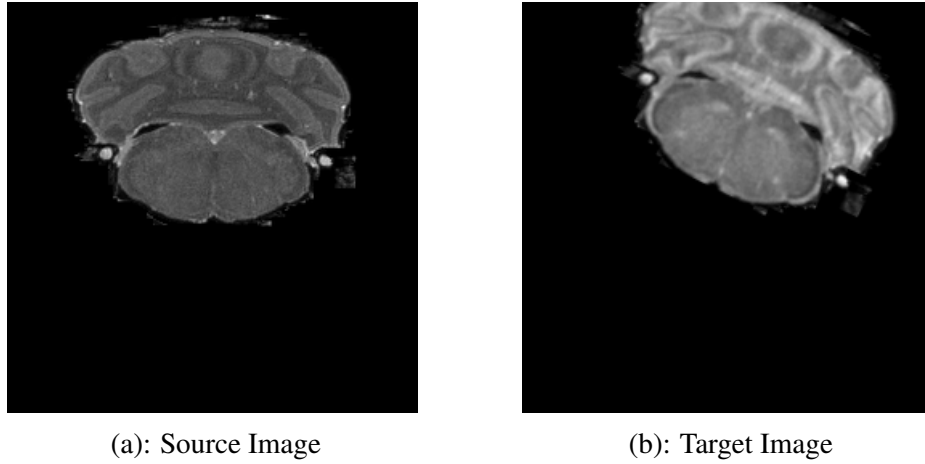
Figure 4.20: Mouse brain image slice 120

1. The randomly generated sample is partitioned into two subsamples:

   - Training set (size: 1000 along with the corresponding estimations of Mutual information (MI))

   - Test set (size: 500 along with the corresponding estimations of Mutual information (MI))

2. Construction of the approximation $\widetilde{MI}$ using the Training set.

3. Computation of the Mean Absolute Error of the i-th partition (MAE_i) using the

$$MAE_i = \frac{1}{N} \sum_{j=1}^{N} \left| \widetilde{MI}(x_j) - MI(x_j) \right| \qquad (4.2)$$

where N is the size of the Test set and x_j is the j-th transformation.

These steps are for one partition. In the experiments, the initial set is partitioned randomly twenty times. Therefore, the aforementioned steps are executed twenty times. In the end, the Mean Absolute Error (MAE) is calculated:

$$MAE = \frac{1}{20} \sum_{i=1}^{20} MEA_i \qquad (4.3)$$

The Mean Absolute Error is calculated for each method and image pair. Indicative images are presented in Fig 4.40-4.23, where the right subfigure is a randomly transformed version of the left subfigure.

(a): Source Image   (b): Target Image

Figure 4.21: 78 Greece Fires

(a): Source Image   (b): Target Image

Figure 4.22: Shrinking Meredith Lake

<div align="center">

(a): Source Image   (b): Target Image

Figure 4.23: 81 Oahe Reservoir

</div>

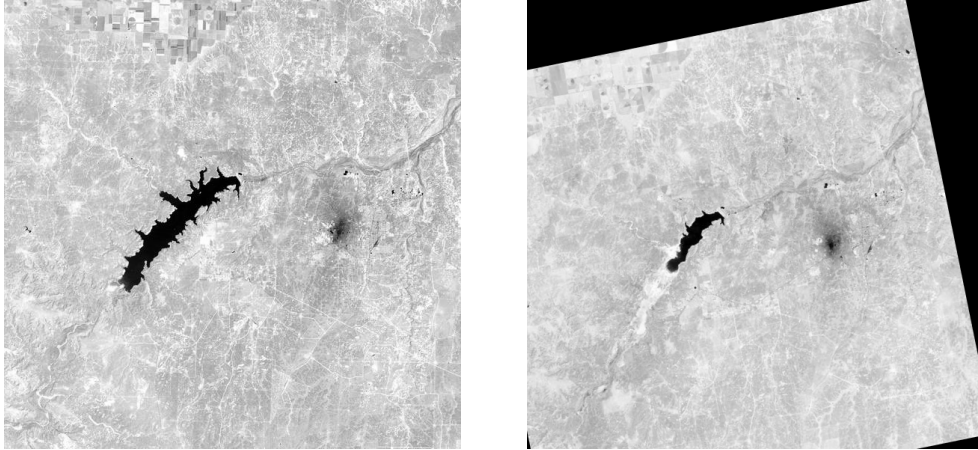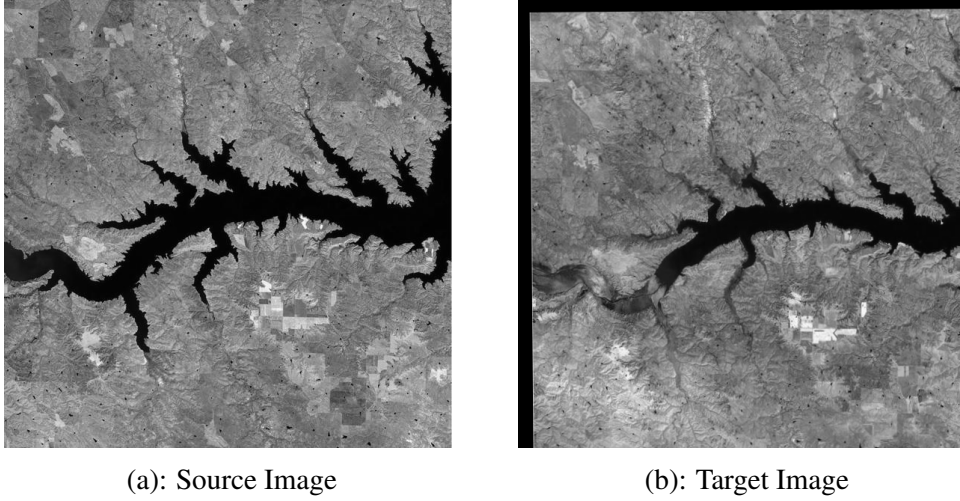In Fig.4.24, the MEA for each image pair and approximation method is presented. MLP has the worst performace of all. Although rbf network performs far better than MLP, it is no match for krr and krls. Eventually, it is SVR that produces the minimum MAE. Also, in Fig.4.25, the mean duration of a single similarity estimation using SVR is presented and what is observed is its small duration which lies in the range of $(5 \cdot 10^{-6}, 3.5 \cdot 10^{-5})$ seconds. In the next experiments, SVR was inserted in the new HA variant and new experiments were conducted in order to see how the computational cost is reduced. The variant, which uses the SVR approximation method, is called EHAR, while the original variant is called HAR. where, Tr_s is the training set and s_i, $i = 1 \ldots, 5$ is the size of Tr_s. Initially, an empty training sample is created, in which the initial population of random solutions is inserted. Until the size of the training set becomes equal to s_1, the original mutual information is used for the evaluation of the new solutions, which are then inserted into the training set. When its size is s_1, an approximation using SVR is constructed and then an alternation between the original Mutual Information and its approximation begins. In every 2*POPSIZE new solutions, (POPSIZE-1) of them solutions are evaluated using original Mutual Information and are included in the training set and then (POPSIZE +1) new solutions are evaluated using the approximation. When the training sample's size becomes equal to s_2, s_3, s_4 and s_5, new and more accurate (because of the larger size) approximations are constructed. The search stops when the maximum number of iterations is reached or the candidate solutions have reached a plateau where any more search is pointless. In the registration experiments, the images are the same as the ones in the previous experiments of determining the best approximation method. Also, EHAR was compared with HAR and two other methods from ITK, which use the following optimization methods:

- Regular Step Gradient Descent

- One Plus One Evolution Strategy

For each image pair, each optimization method was used in 100 image registration experiments to test the ability of each method to find the global optimum as well as the computational cost. In the Table **??**, we see the optimization methods results. The numbers in each cell of the matrix are the successful experiments for each method and image pair.
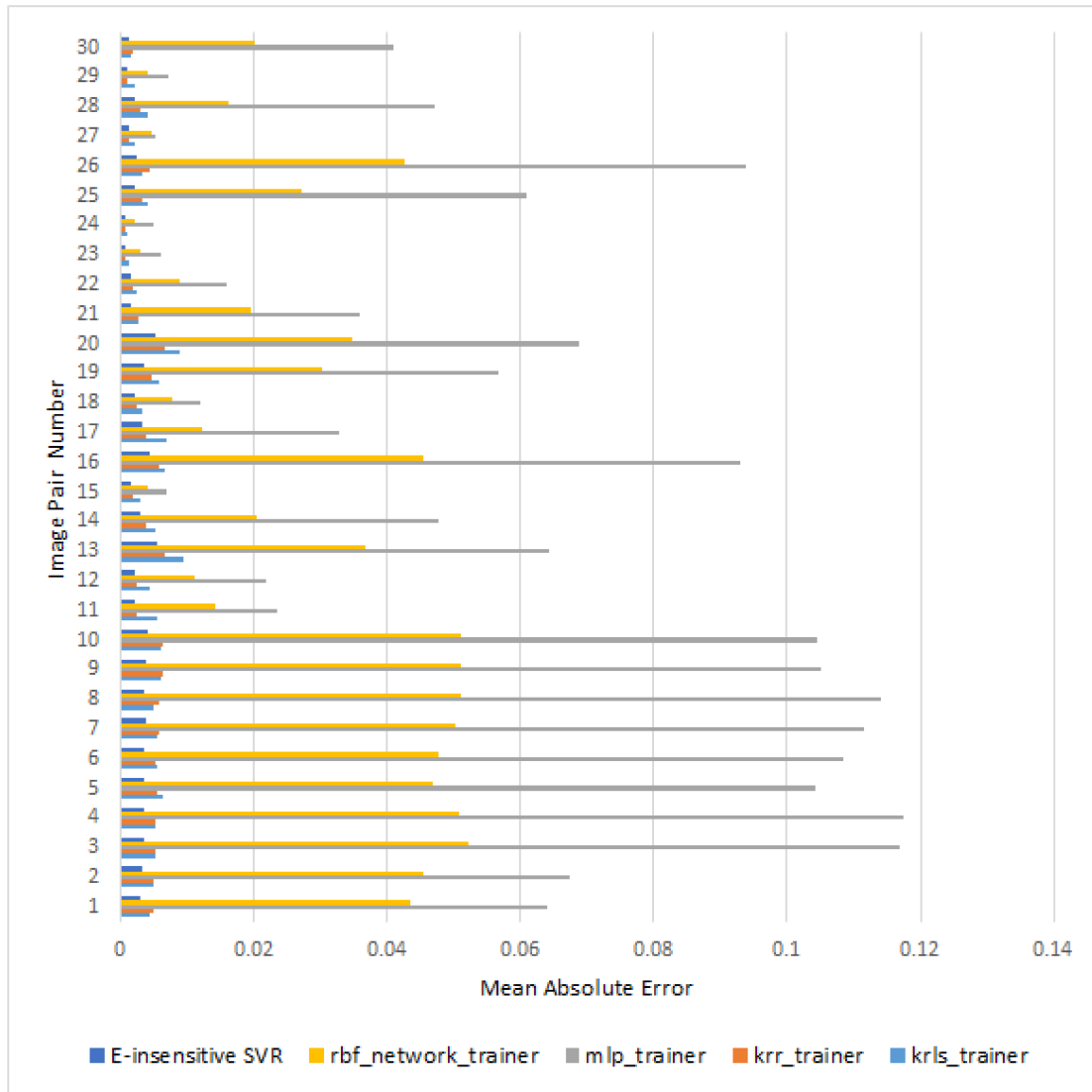
<div align="center">

84

</div>

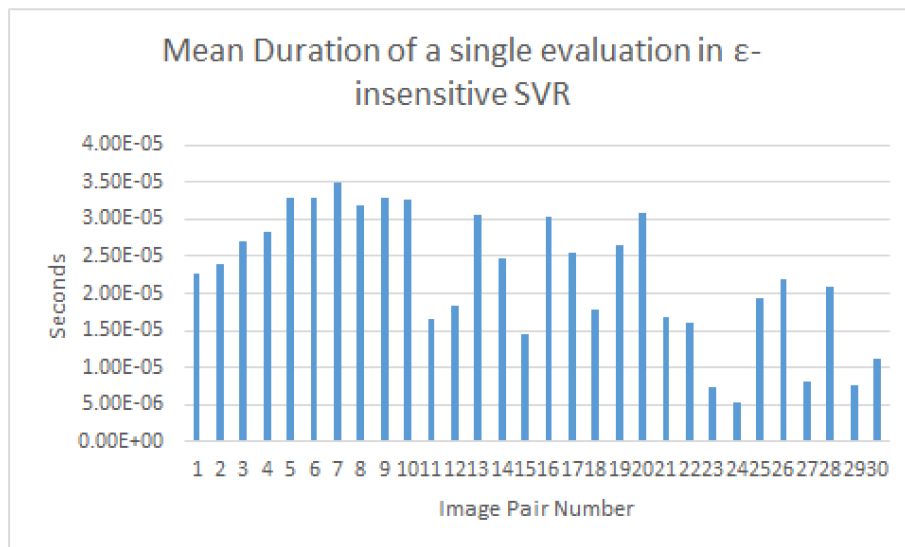Figure 4.24: MEA for each image pair and approximation method

Figure 4.25: Mean Duration of a single similarity estimation using SVR

Data: h_{min},h_{max},par_{min},par_{max},$step\_size$,POPSIZE,MAX_ITER, s_i, $i = 1\ldots,5$

Result: best individual

1   Create initial training set $Tr_s = \emptyset$;

2   Create initial candidate population P of size(**POPSIZE**);

3   $Tr_s = Tr_s \cup P$ **for** *i=1:MAX_ITER* **do**

4      $h = h_{min} + (i/MAX\_ITER) \cdot (h_{max} - h_{min})$;

5      $par = par_{min} + (i/MAX\_ITER) \cdot (par_{max} - par_{min})$;

6   Create $new\_harmony$;

7   **for** *j=1:NVAR* **do**

8   **if** $U(0,1) < h$ **then**

9   Choose $k = U\_INT(1, MAX\_ITER)$;

10            $new\_harmony$(j)=$P[k][j]$;

11   **if** $U(0,1) < par$ **then**

12   Choose $d = U\_INT(1, MAX\_ITER)$, $d \neq k$;

13               $new\_harmony$(j)+=$best(j) + c(best(j) - P[k][j])\frac{(fun(best)-fun(P[k]))}{abs(fun(P[k])-fun(P[d]))} + distribution(0,\delta)$;

14   **end**

15   **end**

16   **else**

17            $new\_harmony$(j)=$U(min_j, max_j)$;

18   **end**

19   **end**

20   **if** $|Tr_s| > s_1$ *and i%POPSIZE<POPSIZE/2* **then**

21   Use SVR approximation;

22   **end**

23   **else**

24   Use original MI;

25         $Tr_s = Tr_s \cup new\_harmony$

26   **end**

27   **if** $|Tr_s| = s_i$, $i = 1,\ldots,5$ **then**

28   Create new SVR approximation

29   **end**

30   **if** *new_harmony better than the worst harmony of P* **then**

31   Replace the latter with the new_harmony;

32   **end**

33   **end**

34   best individual= the best P;

35   return best individual;

Algorithm 18: EHAR Algorithm
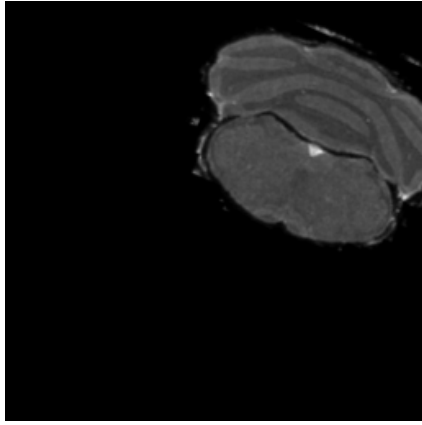
(a): Result with RGD
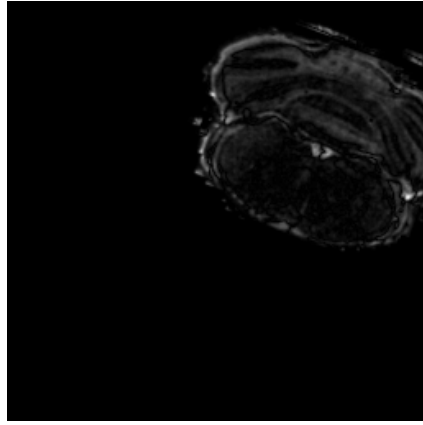
(b): Difference image with RGD

(c): Result with (1+1)ES

(d): Difference image with (1+1)ES

(e): Result with Harmony Search (without SVR)

(f): Difference image with Harmony Search (without SVR)

(g): Result with Harmony Search (with SVR)

(h): Difference image with Harmony Search (with SVR)

Figure 4.26: Image Registration of MRPD and MRT2 at Pelvis area

(a): Result with RGD

(b): Difference image with RGD

(c): Result with (1+1)ES

(d): Difference image with (1+1)ES

(e): Result with Harmony Search (without SVR)

(f): Difference image with Harmony Search (without SVR)

(g): Result with Harmony Search (with SVR)

(h): Difference image with Harmony Search (with SVR)

Figure 4.27: Image Registration of MRPD and MRT2 at Pelvis area
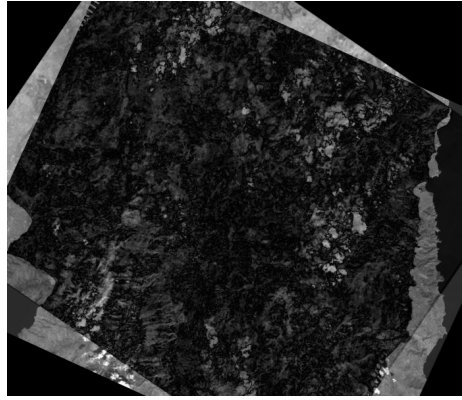
(a): Result with RGD

(b): Difference image with RGD
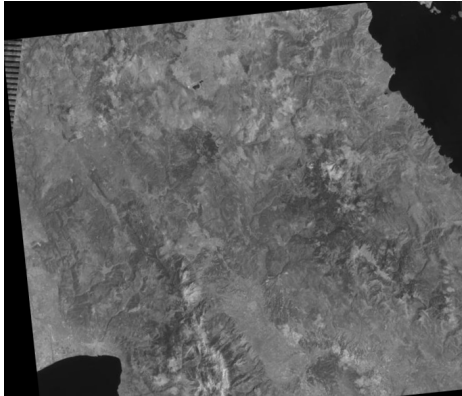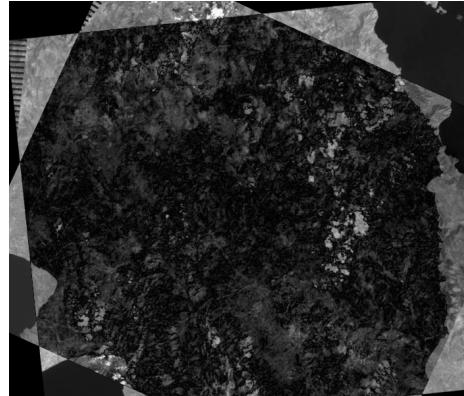
(c): Result with (1+1)ES

(d): Difference image with (1+1)ES
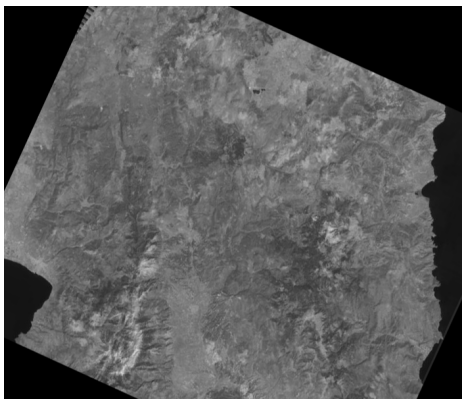
(e): Result with Harmony Search (without SVR)
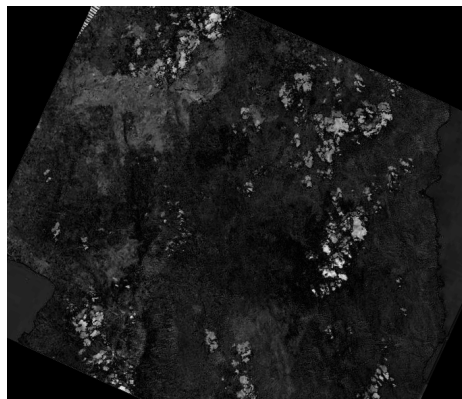
(f): Difference image with Harmony Search (without SVR)
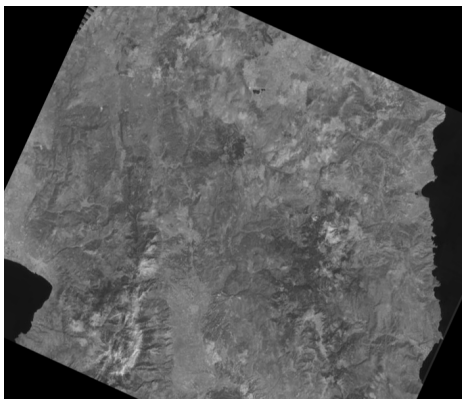
(g): Result with Harmony Search (with SVR)
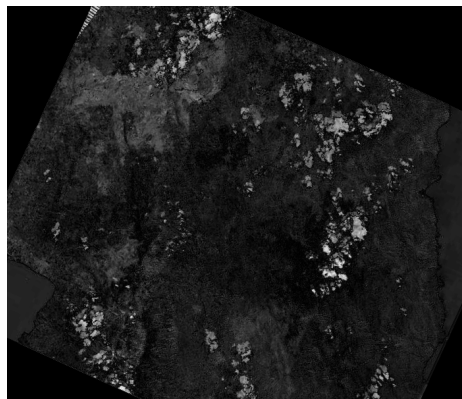
(h): Difference image with Harmony Search (with SVR)

Figure 4.28: Image Registration of MRPD and MRT2 at Pelvis area
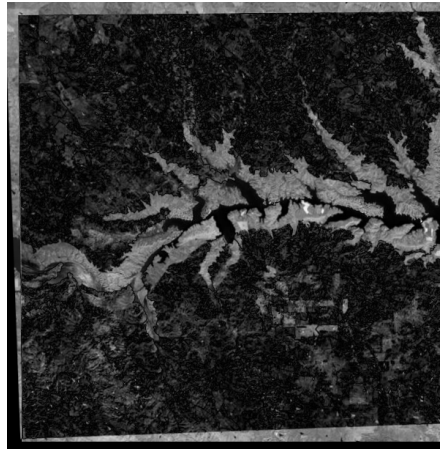
(a): Result with rgd

(b): Difference image with rgd

(c): Result with (1+1)ES

(d): Difference image with (1+1)ES

(e): Result with Harmony Search (without SVR)

(f): Difference image with Harmony Search (without SVR)

(g): Result with Harmony Search (with SVR)

(h): Difference image with Harmony Search (with SVR)

Figure 4.29: Image Registration of MRPD and MRT2 at Pelvis area
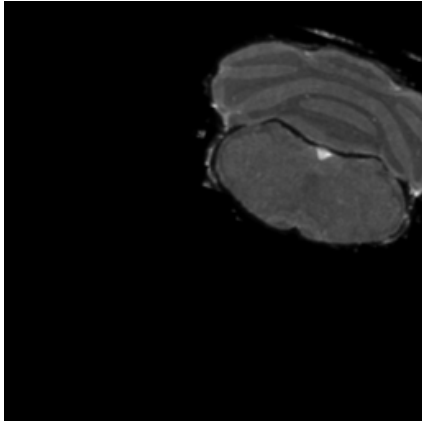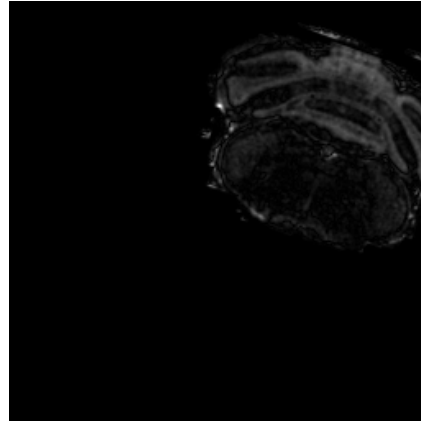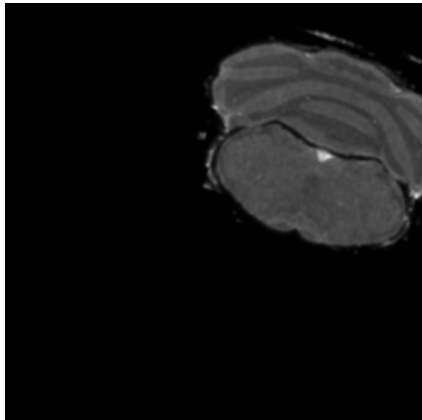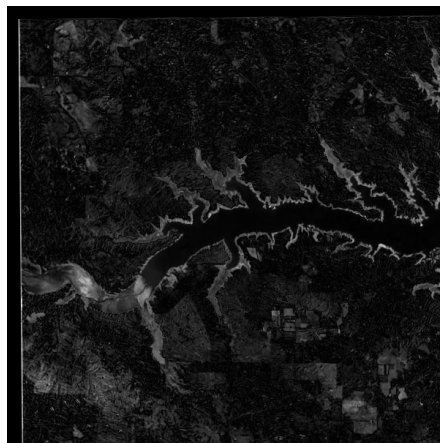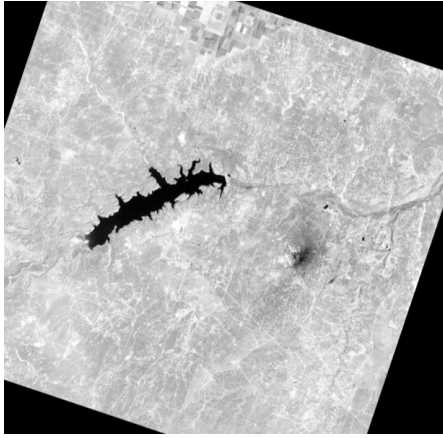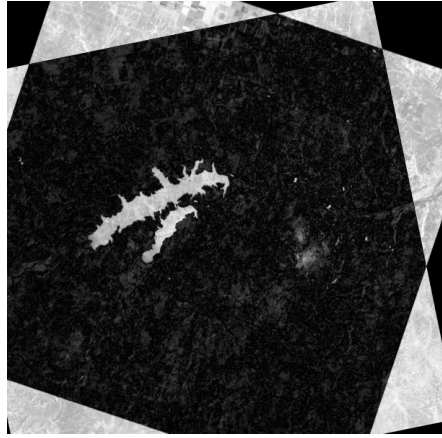
(a): Result with rgd
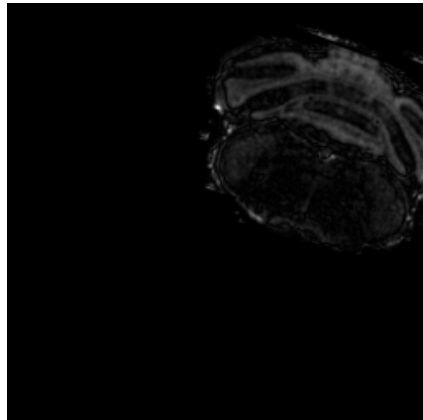
(b): Difference image with rgd
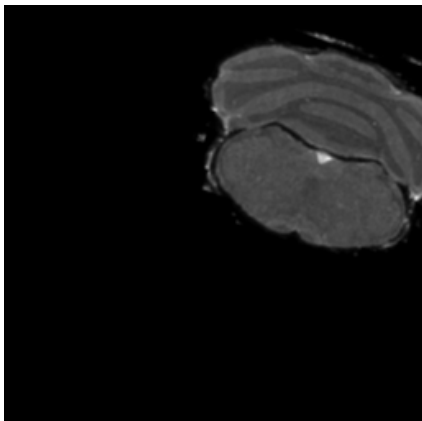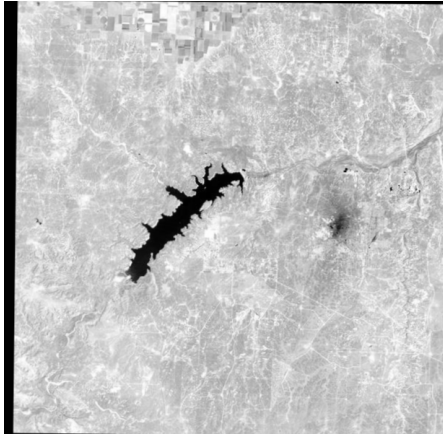
(c): Result with (1+1)ES
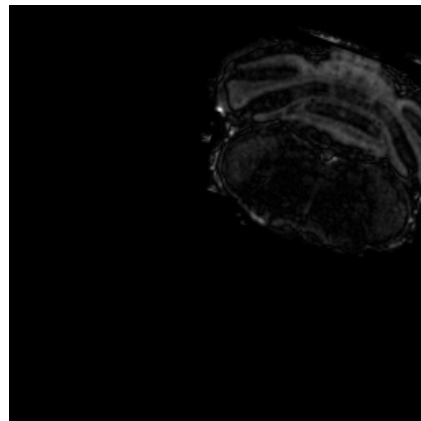
(d): Difference image with (1+1)ES

(e): Result with Harmony Search (without SVR)

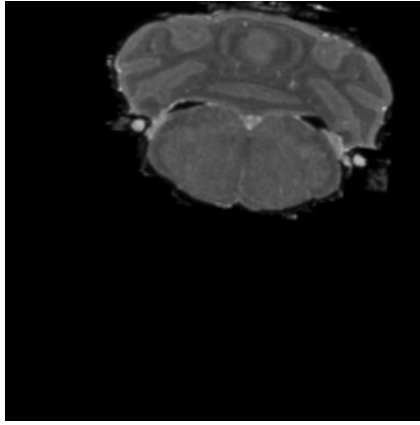(f): Difference image with Harmony Search (without SVR)

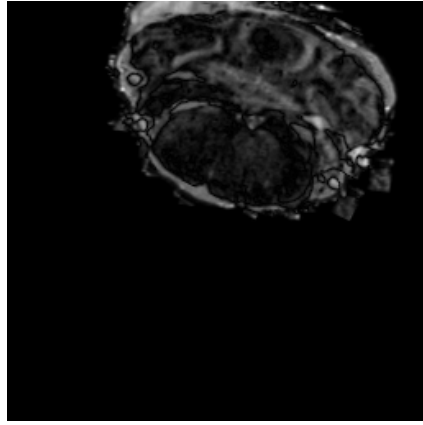(g): Result with Harmony Search (with SVR)

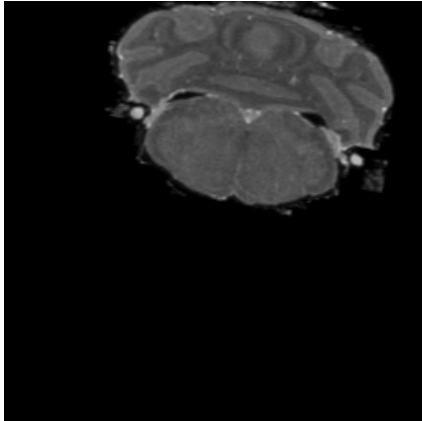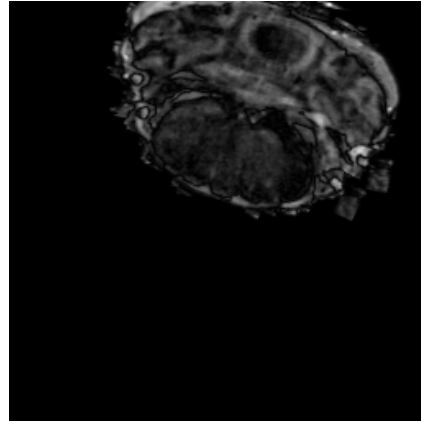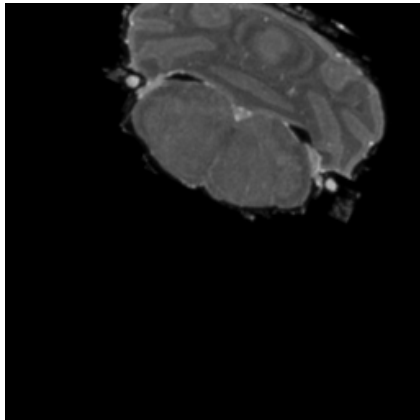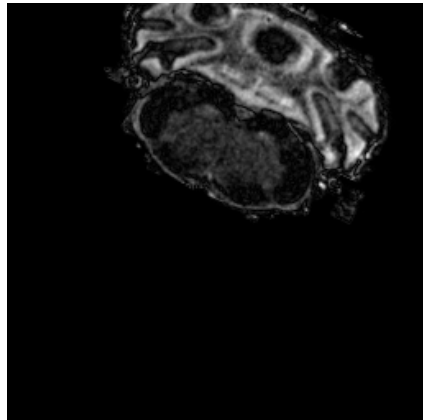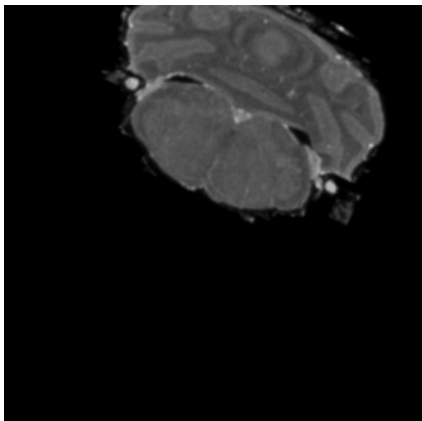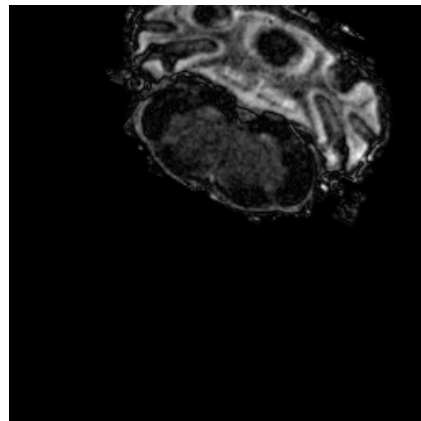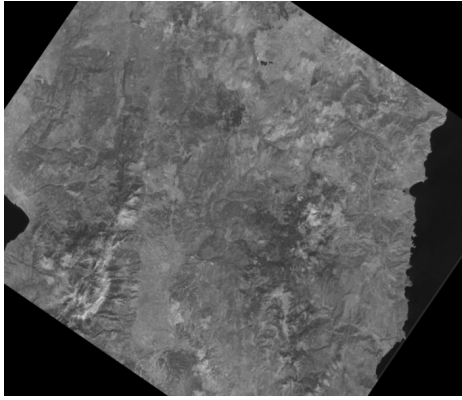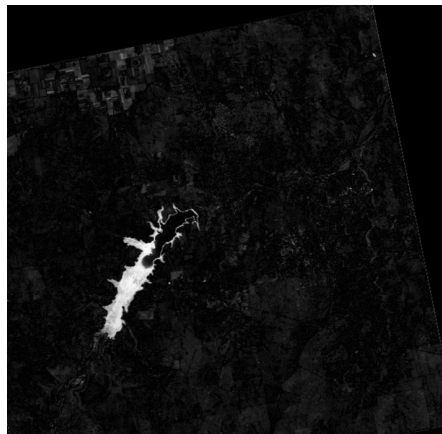(h): Difference image with Harmony Search (with SVR)

Figure 4.30: Image Registration of MRPD and MRT2 at Pelvis area

Table 4.4: Number of successful experiments

No	Experiment	RSGD	(1+1)ES	HAR	EHAR
1	Mouse T2 110	50	99	100	100
2	Mouse T2 120	46	60	100	100
3	Mouse T2 130	65	72	100	100
4	Mouse T2 140	72	100	100	100
5	Mouse T2 150	38	95	100	100
6	Mouse T2 160	69	100	100	100
7	Mouse T2 170	13	100	100	100
8	Mouse T2 180	16	100	100	100
9	Mouse T2 190	70	100	100	100
10	Mouse T2 200	89	100	100	100
11	I01	10	34	100	100
12	I02	5	100	100	100
13	I03	7	100	100	100
14	I04	55	100	100	100
15	I05	5	0	100	100
16	I06	3	49	100	100
17	I07	0	100	100	100
18	I08	0	94	100	100
19	I09	10	98	100	100
20	I10	0	100	100	100
21	78Greece_Fires	6	0	100	100
22	81Oahe	20	100	100	100
23	210shrinking	0	37	100	100
24	Aguascalientes	0	0	100	100
25	Bangong	64	84	100	100
26	Gibraltar	28	100	100	100
27	Greenland_Fires	0	0	100	100
28	Lake_Meadwater	31	75	100	100
29	Lake_Shasta	0	100	100	100
30	Lake_Urmia	45	100	100	100

Indicative results of the image registration experiments are the Figures 4.26-4.30. In detail, the left subfigures depict the results of the image registration using RGD,(1+1)ES and the Harmony Search variant with and without the SVR-based Surrogate model, while the right subfigures show the how close the result is to the reference image.. It is obvious that in all cases the Harmony Search variant (with and without the Surrogate model) manages to align the images. In Table 4.4 , examples, where EHAR and HAR succeeded and RGSD and $(1 + 1)$ ES failed, are presented. The following are observed:

- RSGD has the worst performance of all, due to the fact that it has no randomness in order to escape local optima.

- One-Plus-One can deal much better than RSGD, because the randomness that is introduced into the algorithm can help it escape local optima. However, since it optimizes only one solution, it has limited information of the search space and, therefore, is also prone to entrapment in local minima

- Both HAR and EHAR succeed in each image registration experiment for each image pair. However, EHAR is less computationally expensive as we see in
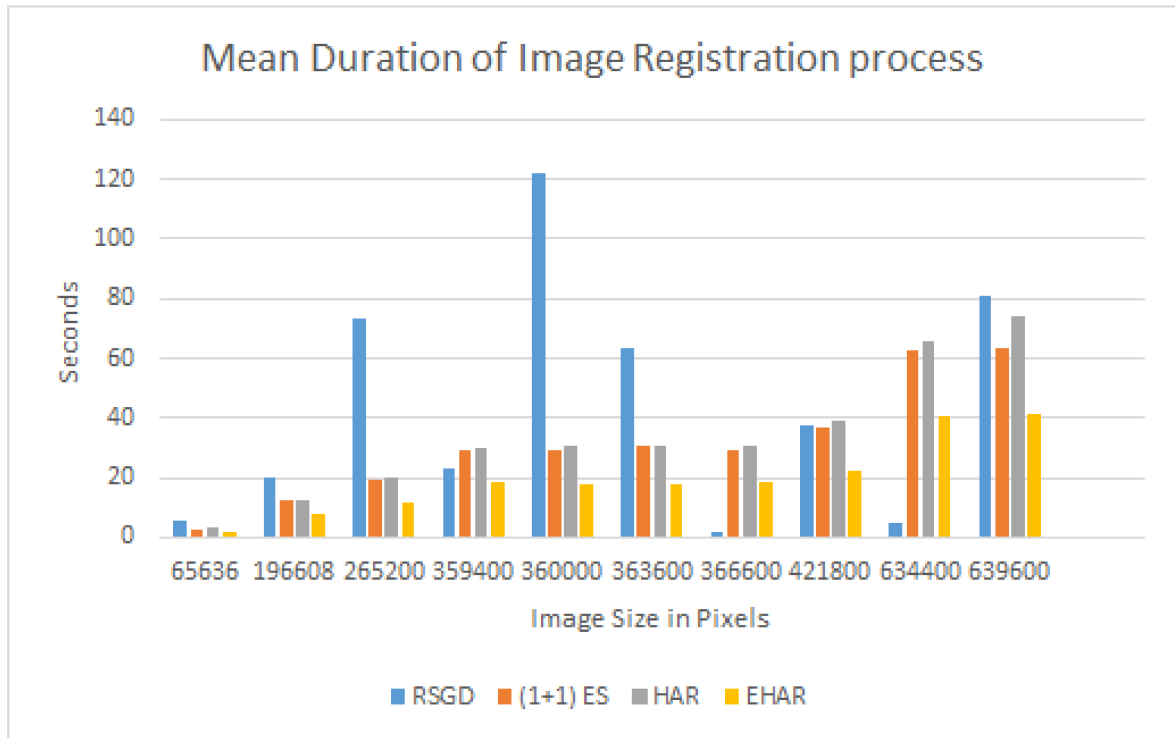
Figure 4.31: Mean duration of Image registration process with respect to size
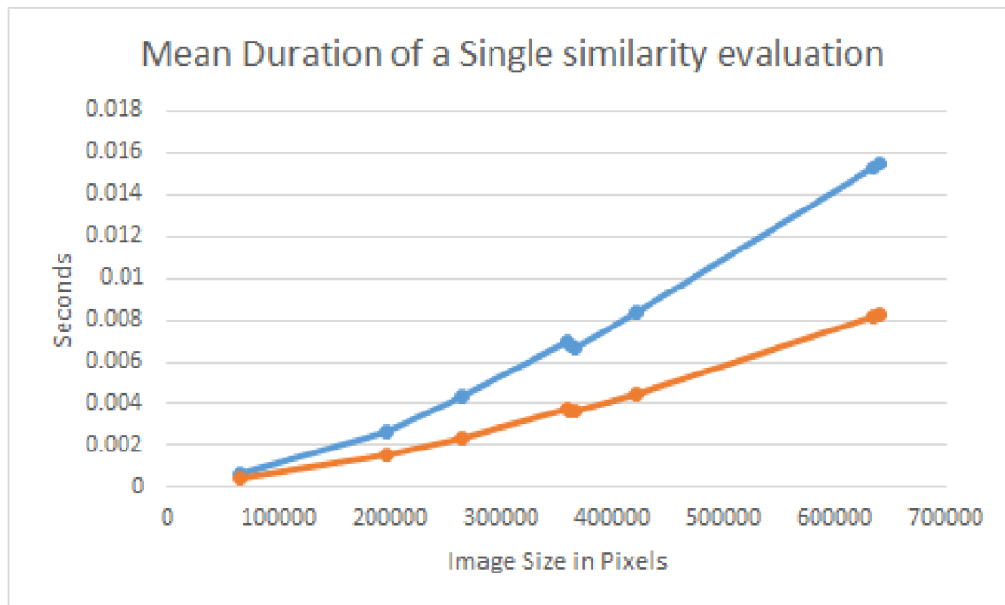
Figure 4.32: Mean duration of Similarity estimation process with respect to size

Fig. 4.32, because of the partial substitution of the original Mutual Information by it SVR approach, which, because of its cheap computation, lowers significantly the duration as we see in Fig.4.18. In fact, the reduction of the duration is up to 47%. The impact of the SVR approach increases as the the image size increases.

4.1.4 Image Stitching

As it is aforementioned, Image Registration is the first and most important step in Image Stitching and must have the following characteristics:

- Robustness and Accuracy

- Capable of dealing with any image pair.

In our experiments, we came with the following problems:

- Increased Image size (up to 2560×1920)

- Increased Search Space

These problems made the need for a method that is capable of exploring a vast space and produce results in the minimum computational cost. Our conrtibution is a new optimization method that combines the following:

- A new, more robust optimization method which is an enhanced variant of Harmony search.

- Subsampling via downscaling the images. The reasons for using this kind of subsampling are the following ones:

  - Primal Reduction of computational cost
  - Reduction of the size of the image leads to reduction of the translation range and, therefore, the need for further search.

- Use of Support Vector Regression for the reduction of the computational cost.

(a): First Row

(b): Second Row

(c): Third Row

(d): Fourth Row

(e): Fifth Row
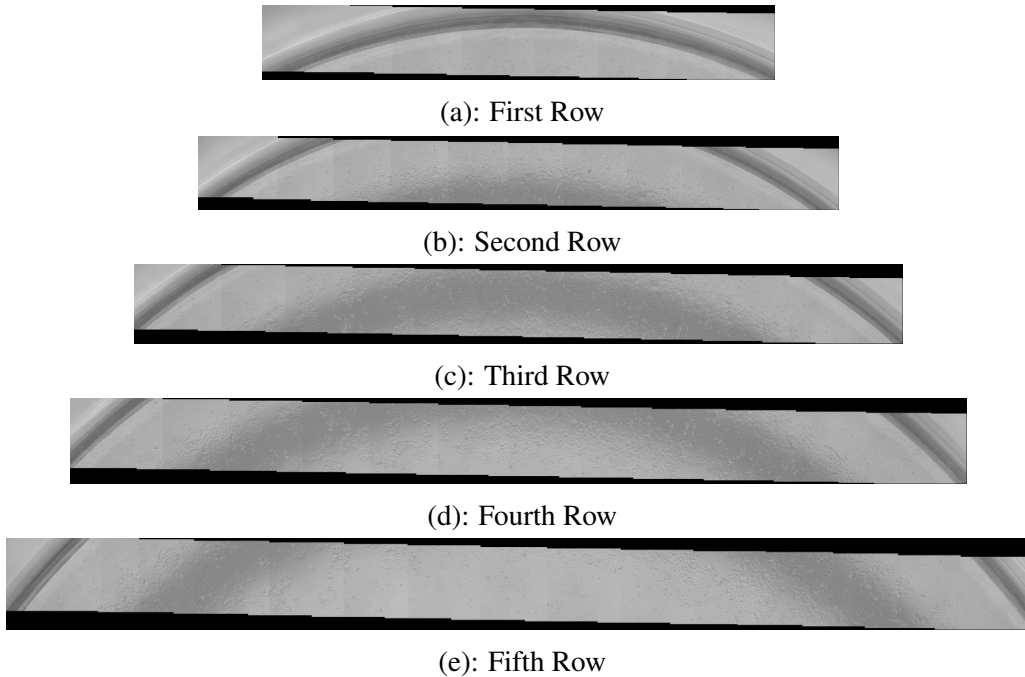
Figure 4.33: Results of Row Stitching
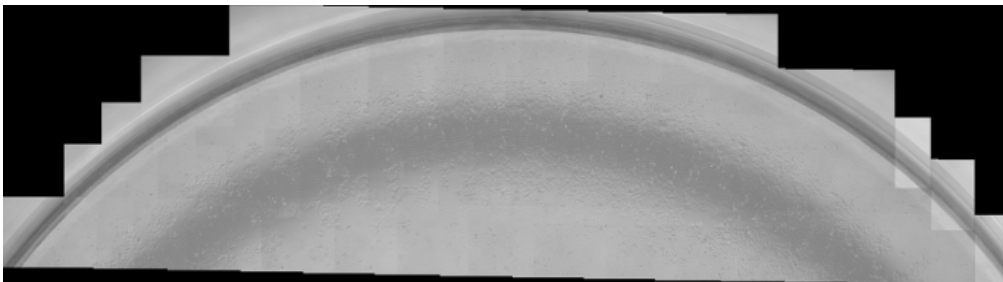
Figure 4.34: Result of first example

In Fig. 4.33,4.34, the successful results of stitching the images per row and the final stitching of the previous stitched rows are presented. It is the example where the FIJI method failed to stitch the images properly.
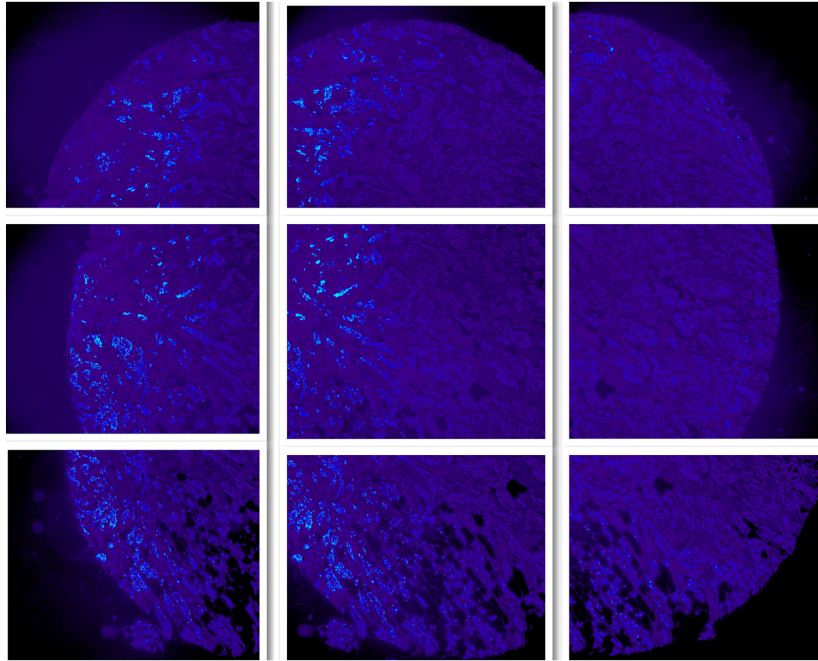
(a): Set of Microscopic Images
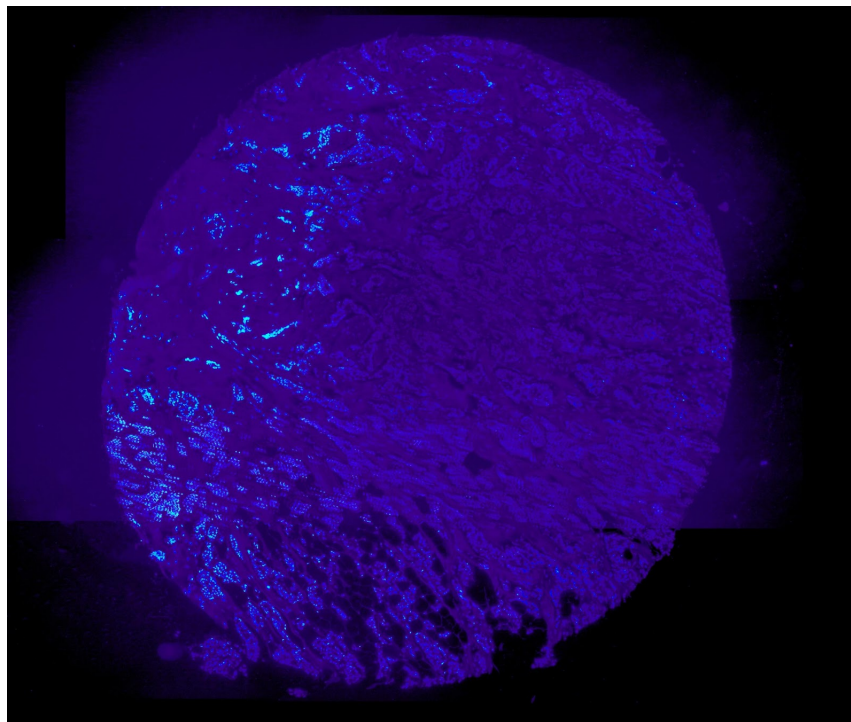
(b): Final result

Figure 4.35: Second example of Image Stitching
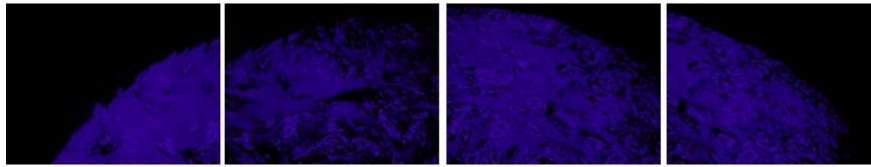
(a): First Row
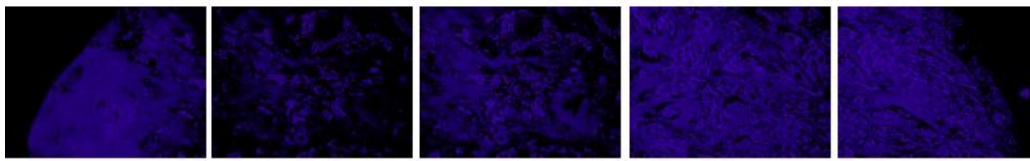
(b): Second Row
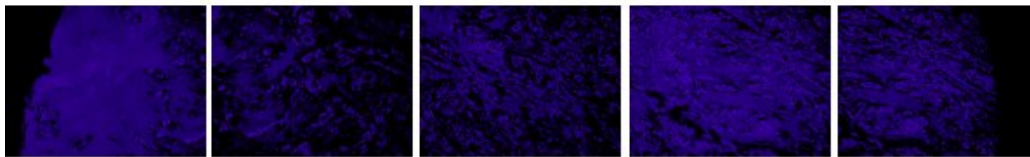
(c): Third Row
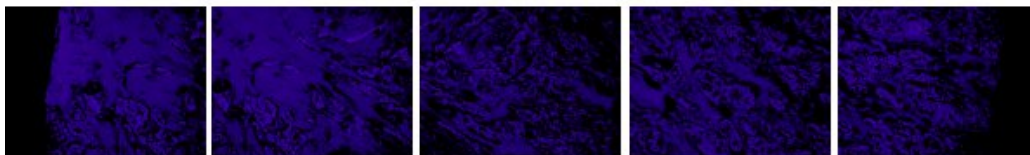
(d): Fourth Row
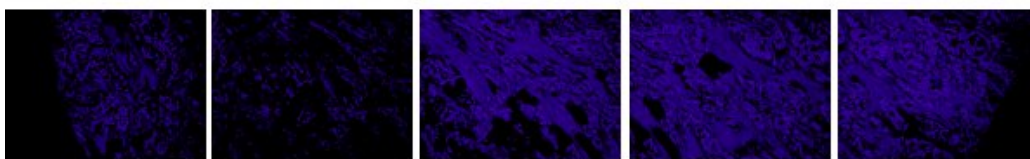
(e): Fifth Row
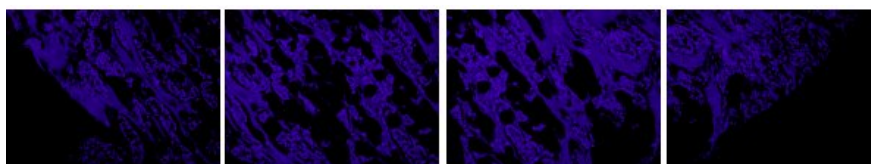
(f): Sixth Row
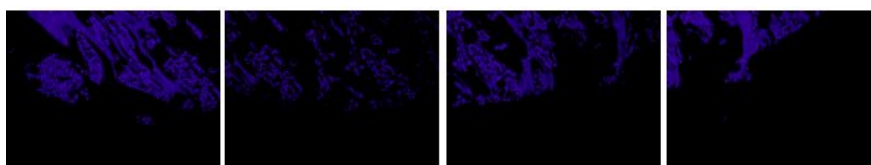
(g): Seventh Row

Figure 4.36: Set of Microscopic Images
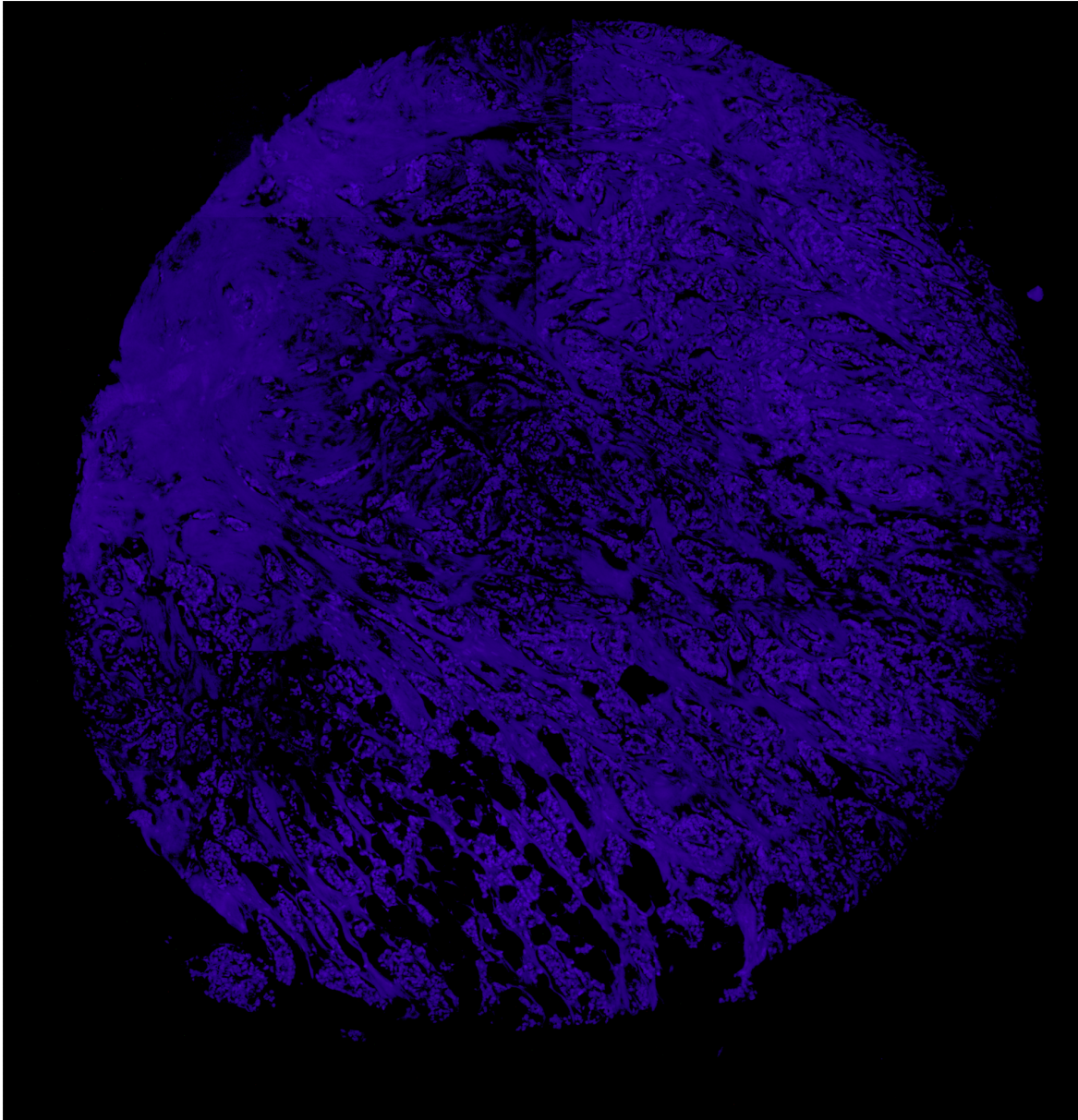
Figure 4.37: Final Result

In Fig. 4.35, a second successful example of image stitching is presented, where the acquired images are form a rectangular grid. A third example is presented in Fig. 4.36, where the images do not form a rectangular grid. These images are stitched successfully with my method in Fig. 4.37.
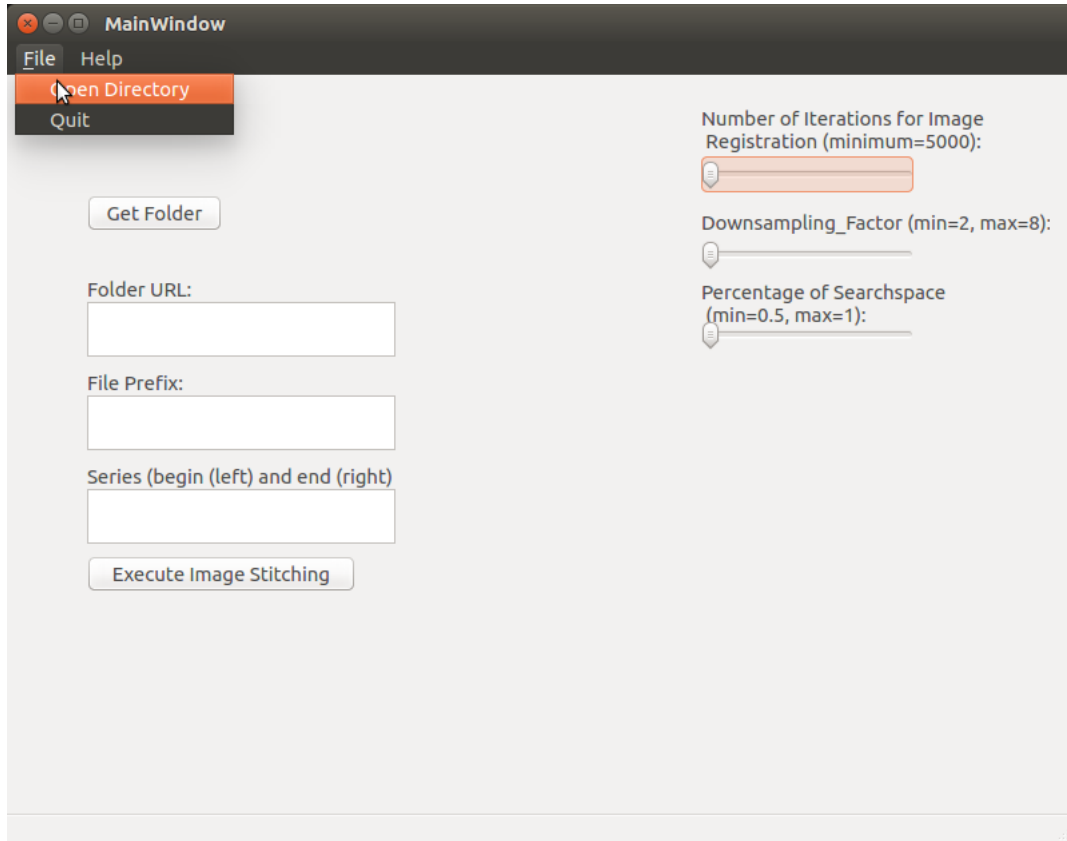
Figure 4.38: GUI Interface

4.2 Image Comparison Using Renyi Divergence

In the presented Tables I-III, we observe that the accuracy regarding rotation, x translation and y translation, increases on average along with the a parameter value. Similarly, increase of the sample size leads (on average) to more accurate results (albeit obtained with greater computational cost). In Tables IV and V, we present the Mean Absolute Error in Rotation and Translation with respect to the α parameter and the subsampling factor respectively. In Table IV, we see that the increase of the α parameter reduces the MAE. The same applies for the percentage of the pixels used in subsampling, as shown in Table V.

Table 4.5: Error in Rotation

	Renyi Parameter							
%	0.1	0.3	0.5	0.7	0.9	1	1.1	1.3
1	9.81E-03	8.48E-03	8.59E-03	8.51E-03	7.97E-03	7.78E-03	7.60E-03	7.28E-03
2	6.29E-03	6.01E-03	5.95E-03	5.81E-03	5.75E-03	5.67E-03	5.83E-03	5.60E-03
4	5.21E-03	4.73E-03	4.53E-03	4.51E-03	4.35E-03	4.46E-03	4.41E-03	4.53E-03
6	4.89E-03	3.86E-03	4.00E-03	3.99E-03	3.77E-03	3.66E-03	3.96E-03	3.76E-03
8	4.41E-03	3.83E-03	3.55E-03	3.38E-03	3.55E-03	3.43E-03	3.34E-03	3.47E-03
10	4.13E-03	3.67E-03	3.29E-03	3.21E-03	3.31E-03	3.16E-03	3.34E-03	3.09E-03
20	3.96E-03	3.00E-03	2.61E-03	2.68E-03	2.66E-03	2.66E-03	2.54E-03	2.67E-03

Table 4.6: Error in X Translation

%	Renyi Parameter							
	0.1	0.3	0.5	0.7	0.9	1	1.1	1.3
1	1.19E+00	1.26E+00	1.17E+00	1.13E+00	1.08E+00	1.06E+00	1.03E+00	1.02E+00
2	8.42E-01	7.76E-01	7.77E-01	7.69E-01	7.52E-01	7.75E-01	7.50E-01	7.38E-01
4	6.13E-01	5.73E-01	5.67E-01	5.68E-01	5.66E-01	5.46E-01	5.50E-01	5.57E-01
6	5.65E-01	4.85E-01	4.72E-01	4.64E-01	4.71E-01	4.60E-01	4.59E-01	4.86E-01
8	4.95E-01	4.50E-01	4.18E-01	4.33E-01	4.28E-01	4.22E-01	4.21E-01	4.25E-01
10	5.19E-01	4.39E-01	3.90E-01	3.72E-01	3.81E-01	3.69E-01	3.68E-01	3.71E-01
20	4.77E-01	3.59E-01	3.08E-01	3.14E-01	3.00E-01	2.95E-01	2.92E-01	3.07E-01

Table 4.7: Error in Y Translation

%	Renyi Parameter							
	0.1	0.3	0.5	0.7	0.9	1	1.1	1.3
1	1.15E+00	1.17E+00	1.17E+00	1.10E+00	1.02E+00	1.02E+00	1.06E+00	9.90E-01
2	8.53E-01	7.99E-01	7.43E-01	7.68E-01	7.87E-01	7.59E-01	7.21E-01	7.48E-01
4	6.00E-01	5.82E-01	5.61E-01	5.40E-01	5.43E-01	5.36E-01	5.61E-01	5.58E-01
6	5.72E-01	5.15E-01	4.77E-01	4.69E-01	4.67E-01	4.57E-01	4.73E-01	4.76E-01
8	5.26E-01	4.61E-01	4.28E-01	4.01E-01	4.33E-01	4.11E-01	4.21E-01	4.35E-01
10	5.17E-01	4.38E-01	3.95E-01	3.80E-01	3.94E-01	4.00E-01	3.89E-01	3.86E-01
20	4.73E-01	3.58E-01	3.12E-01	3.10E-01	2.93E-01	2.86E-01	2.87E-01	3.14E-01

Table 4.8: Average error with respect to the α parameter

Parameter	Renyi Parameter							
	0.1	0.3	0.5	0.7	0.9	1	1.1	1.3
Rotation	5.53E-03	4.80E-03	4.64E-03	4.58E-03	4.48E-03	4.40E-03	4.43E-03	4.34E-03
X Trans	6.72E-01	6.21E-01	5.86E-01	5.79E-01	5.69E-01	5.61E-01	5.53E-01	5.58E-01
Y Trans	6.70E-01	6.17E-01	5.83E-01	5.66E-01	5.63E-01	5.54E-01	5.59E-01	5.58E-01

Table 4.9: Average error with respect to the subsampling factor

Parameter	Percentage						
	1	2	4	6	8	10	20
Rotation	8.25E-03	5.86E-03	4.59E-03	3.99E-03	3.62E-03	3.40E-03	2.85E-03
X Trans	1.12E+00	7.72E-01	5.68E-01	4.83E-01	4.37E-01	4.01E-01	3.32E-01
Y Trans	1.08E+00	7.72E-01	5.60E-01	4.88E-01	4.40E-01	4.13E-01	3.29E-01

4.3 Affine Registration

In this section, we present the results of the affine image registration experiments, where Genetic Algorithms, the first Harmony Search and the second Harmony Search (the last one is tested with and without the surrogate modeling based on Support Vector Regression) variant are compared. In the experiments of Affine Registration, the images are acquired from the brain and are of two modalities:

- MRPD

- MRT1

In these images, noise was added in order to test the optimization capabilities of the Elitist GA, the first and second Harmony Search Variants. For the experiments of the registration, each optimization algorithm had the following parameters:

- Elitist GA:

    - Population Size:50
    - Elites:5
    - Number of Generations: 1600
    - Cross rate:0.7
    - Mutation rate: 0.3

- Harmony Search Variants Parameters:

    - Population Size:50
    - h: minimum=0.0, maximum=0.7
    - par: minimum=0.0, maximum=0.3
    - Number of Iterations (first HS variant): 80000
    - Number of Sessions (second HS variant): 1600

The images were downloaded from ..., and random rigid and scaling transformations were applied on the images. The range of the transformations were:

- Scaling factor: $[0.8, 1.2]$

- Rotation: $[-0.50, 0.50]$ radians

- Translation along axes x, y: $[-50, 50]$ pixels

The aforementioned optimization methods were tested 20 times for each pair of images in order to test their robustness in affine image registration. The result of the image registration of a given pair is compared with the original random image transformation. Then, the Mean Average Error (MAE) is calculated for each parameter of the affine transformation

$$MEA_{i,j} = \frac{1}{20} \sum_{j=1}^{20} |p_{i,j} - p_j| \tag{4.4}$$

where p_j is the $j-$th parameter of the known affine image transformation ($j = 1, \ldots, 4$ are the parameters for the combined rotation, scaling and stretching and $j = 5, 6$ are the parameters for translation along axes x, y), and $p_{i,j}$ is the $j-$th parameter of the result of the $i-$th experiment. In Tables 4.10-4.13, we present the Mean Average Errors for each parameter and image pair using the aforementioned optimization methods.

In Tables 4.10,4.11, the bold-fonted rows represent the image pairs with misregistration errors. In Fig., indicative examples of affine image registration are presented. In both Figures, Elitist GA fails to align the images, while the First HS variant fails in the second Figure. The second HS Variant, both with and without the SVR-based surrogate model, manages to align the images. Finally in Table 4.14 the number of successful results for each image pair and method are presented, while in Table 4.15 the duration of the affine image registration process with and without the surrogate model for each image pair is presented.

Table 4.10: Mean Average Error using GA

Pair	Parameters					
	P1	**P2**	**P3**	**P4**	**P5**	**P6**
1	1.46E-04	2.89E-04	3.00E-04	4.51E-04	5.09E-01	2.83E-01
2	7.82E-05	1.49E-04	1.89E-04	1.17E-04	2.84E-01	2.60E-01
3	**5.54E-02**	**1.22E-02**	**1.95E-01**	**2.01E-02**	**7.66E+00**	**1.27E+01**
4	**5.63E-02**	**2.96E-02**	**1.23E-01**	**4.20E-02**	**5.03E+00**	**1.14E+01**
5	1.50E-04	1.33E-04	3.06E-04	1.65E-04	1.08E-01	6.63E-02
6	7.21E-05	1.74E-04	2.41E-04	1.89E-04	1.64E-01	1.42E-01
7	9.39E-05	2.38E-04	5.43E-04	1.57E-04	7.18E-01	1.22E+00
8	6.06E-05	2.06E-04	2.13E-04	3.13E-04	2.40E-01	4.46E-01
9	2.97E-04	3.80E-04	9.43E-04	3.06E-04	1.37E+00	2.94E+00
10	2.29E-04	1.70E-04	4.22E-04	4.00E-04	1.44E-01	1.38E-01

Table 4.11: Mean Average Errors using first HS

Pair	Parameters					
	P1	**P2**	**P3**	**P4**	**P5**	**P6**
1	3.42E-05	2.31E-04	2.51E-04	4.14E-04	4.76E-01	2.60E-01
2	1.04E-04	1.67E-04	9.12E-05	4.95E-05	3.56E-01	1.56E-01
3	**1.27E-02**	**2.40E-03**	**4.95E-02**	**4.05E-03**	**2.38E+00**	**3.09E+00**
4	1.93E-04	1.49E-04	4.24E-04	1.53E-04	2.50E-01	2.75E-01
5	9.01E-05	5.95E-05	4.29E-04	2.21E-04	2.29E-01	7.69E-02
6	1.11E-04	2.78E-04	4.12E-04	1.99E-04	2.82E-01	2.84E-01
7	4.60E-05	9.48E-05	4.01E-04	8.81E-05	5.15E-01	9.23E-01
8	5.44E-05	2.95E-05	5.78E-05	2.81E-04	1.06E-01	3.07E-01
9	1.66E-04	1.41E-04	5.99E-04	2.97E-04	1.76E+00	1.64E+00
10	2.80E-04	2.27E-04	5.42E-04	3.96E-04	1.29E-01	2.16E-01

Table 4.12: Mean Average Errors using second HS variant without SVR

Pair	Parameters					
	P1	**P2**	**P3**	**P4**	**P5**	**P6**
1	7.21E-05	1.07E-04	1.71E-04	3.66E-04	3.73E-01	1.82E-01
2	9.36E-05	1.70E-04	1.11E-04	1.50E-04	3.03E-01	2.97E-01
3	4.87E-05	4.85E-05	1.12E-04	1.18E-04	4.06E-02	1.72E-02
4	1.12E-04	1.09E-04	2.04E-04	1.63E-04	1.42E-01	1.33E-01
5	9.35E-05	6.36E-05	2.47E-04	1.92E-04	9.21E-02	4.78E-02
6	5.71E-05	1.41E-04	2.06E-04	1.28E-04	1.61E-01	1.03E-01
7	3.45E-05	6.03E-05	2.34E-04	1.07E-04	3.63E-01	4.06E-01
8	6.66E-05	1.15E-04	1.15E-04	3.16E-04	1.11E-01	4.04E-01
9	8.54E-05	7.88E-05	3.65E-04	3.45E-04	1.68E+00	1.04E+00
10	1.91E-04	1.82E-04	4.83E-04	5.06E-04	1.72E-01	1.10E-01

Table 4.13: Mean Average Errors using second HS variant with SVR

Pair	Parameters					
	P1	**P2**	**P3**	**P4**	**P5**	**P6**
1	1.95E-04	3.33E-04	3.54E-04	4.40E-04	4.88E-01	4.26E-01
2	7.31E-05	1.53E-04	1.66E-04	1.29E-04	3.15E-01	3.09E-01
3	1.14E-04	9.97E-05	1.92E-04	1.89E-04	1.18E-01	3.82E-02
4	1.09E-04	1.63E-04	3.67E-04	2.23E-04	2.46E-01	1.93E-01
5	2.64E-04	1.32E-04	4.95E-04	2.65E-04	3.34E-01	1.55E-01
6	8.27E-05	1.54E-04	1.65E-04	2.01E-04	1.74E-01	1.53E-01
7	7.43E-05	8.02E-05	2.09E-04	1.53E-04	4.79E-01	5.10E-01
8	1.23E-04	2.12E-04	2.43E-04	4.07E-04	2.48E-01	6.38E-01
9	1.57E-04	1.27E-04	5.34E-04	3.26E-04	2.02E+00	1.46E+00
10	2.92E-04	1.67E-04	5.19E-04	5.72E-04	2.36E-01	1.86E-01

Table 4.14: Successful Results of Affine Image Registration

Image Pair	Optimization Method			
	Elitist GA	**1st HS Variant**	**2nd HS Variant**	**2nd HS Variant+SVR**
1	20	20	20	20
2	20	20	20	20
3	17	19	20	20
4	17	20	20	20
5	20	20	20	20
6	20	20	20	20
7	20	20	20	20
8	20	20	20	20
9	20	20	20	20
10	20	20	20	20

Table 4.15: Duration of affine registration using second Harmony variant with and without SVR-based Surrogate models

Image Pairs	Duration		Reduction
	Second Variant	**Second Variant with SVR**	
Slice 5	312.05	159.10	49.01%
Slice 10	339.32	176.73	47.92%
Slice 15	303.53	149.15	50.86%
Slice 20	307.34	149.64	51.31%
Slice 25	268.57	149.39	44.38%
Slice 30	286.70	151.88	47.02%
Slice 35	296.84	154.54	47.94%
Slice 40	379.86	181.87	52.12%
Slice 45	314.61	156.67	50.20%
Slice 50	278.34	140.32	49.59%
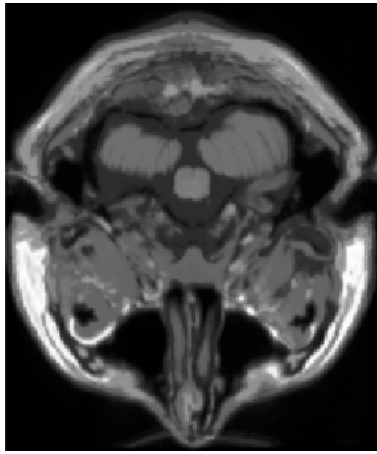
(a): Source Image
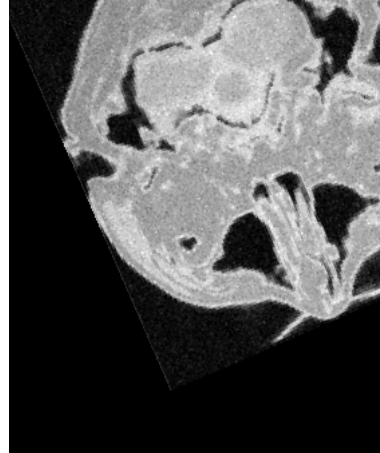
(b): Target Image
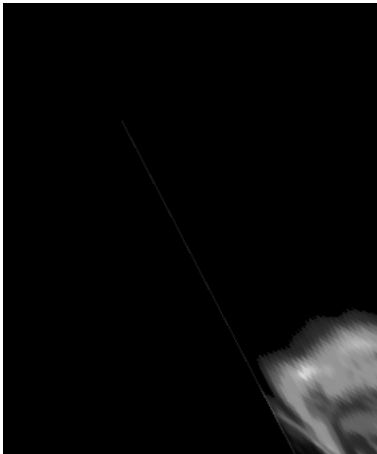
(c): Genetic Algorithm result
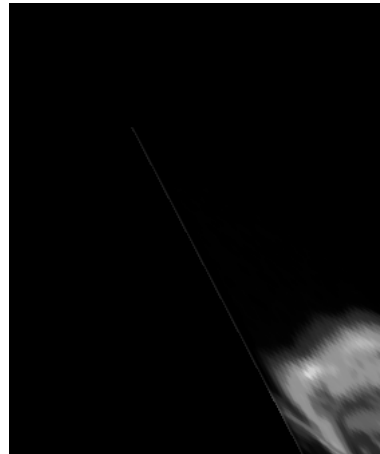
(d): First HS variant result
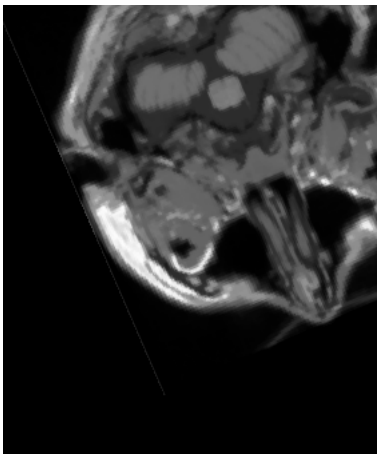
(e): Second HS variant Result
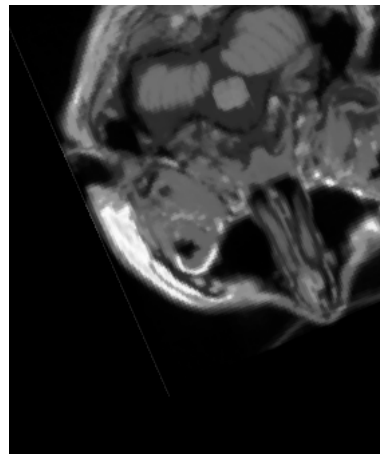
(f): Second HS variant Result with Surrogate model

Figure 4.39: Brain slice 15
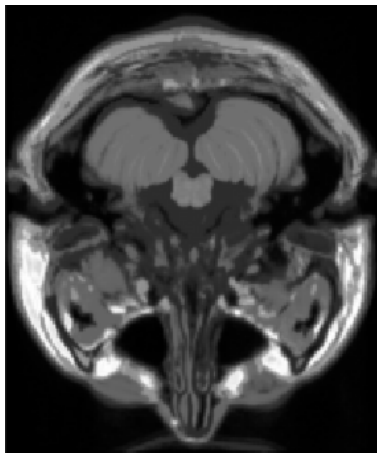
(a): Source Image
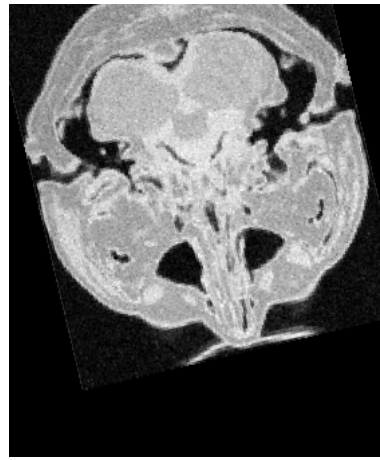
(b): Target Image
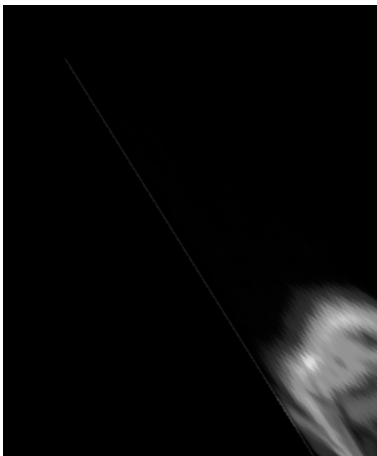
(c): Source Image

(d): Source Image
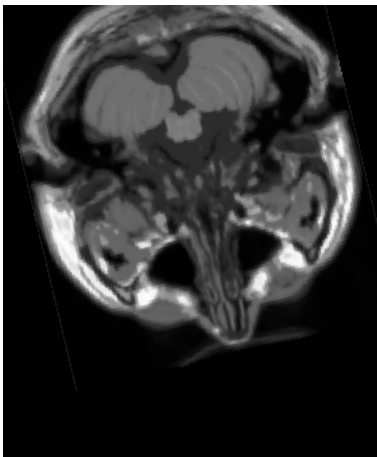
(e): Source Image

(f): Source Image

Figure 4.40: Brain slice 20

Chapter 5

Conclusions and Recommendations

Image registration is a process that is used in many application areas. Depending on, on the one hand, the problems that need to be solved, and, on the other hand, the challenges they may face, many image registration methods have been designed based on the following:

1. Defining the similarity of the images.

2. Defining the appropriate transformation.

3. Defining the appropriate optimization method.

4. Minimization of computational costs

In this dissertation, research was carried out at points 1.3 and 4. The reason why no survey was conducted at point 2 is because, with the exception of the elastic deformations of the images where a large survey has been performed, the basic deformations such as rigid and affine have one single mathematical formula that describes each one of them, therefore there is no ground for research on different ways of modeling these deformations. In the following sections, the conclusions from the above research are presented as well as future plans for applications of image identification as well as for future research.

5.1 Conclusions and insights about the Study

A series of experiments have been executed in order to study the following:

- The performance of elitist genetic algorithms in rigid image registration with respect to the number of elites and the mutation rate

- A way to overcome the disadvantages of Global Harmony Search which is done by designing a new better optimization method.

- A way to overcome the disadvantages faced in image stitching.

- Research in modified mutual-information-like measures based on Renyi divergence and how the parameters of subsampling and Renyi parameters α affect the image registration error

- Introduction of surrogate models in meahuristic-oriented image registration.

- Comparison of Elitist GA and the Harmony Search variants in Affine Image Registration.

In the following subsections, the conclusions as well as the insights are presented.

5.1.1 Elitist GA

In all the experiments of Elitist-GA intensity-based rigid image registration, the optimal transformation (or, at least a good approximation of it) was found, regardless of the number of the elites and the mutation rate. From the images we concluded that as the number of the elites and the mutation rate increases, the mean Mutual Information is increased across all image pairs. That means that on average the elitist method can locate a better approximation of the global optimum. On the other hand, the standard deviation of the Mutual information decreases. That means that there is little variation between the solutions produced in each experiment, therefore it can be safely assume that the probability of diverging from the global optimum is reduced. Considering the above, we conclude that the elitist method can indeed be a very efficient method of for optimization in IIR, as long as the number of the elites and the mutation rate are high. This is an important finding that can add value to several existing rigid registration imaging systems. This is an important finding that can add value to several existing rigid registration imaging systems. For instance, in previous work regarding rigid image registration research using only one elite, we found that elitism surpassed ITK methods in each one of the experiments. Using more elites could therefore increase the accuracy of the results, resulting in increased robustness and decreased registration errors. The idea of preserving the best of the current generation and passing them into the new one can also be exploited into other more sophisticated methods of optimization.

5.1.2 First Harmony Search Variant

As it is analysed, Global Harmony Search has two main disadvantages:

- Fixed values of parameters $h,\ par$

- No use of perturbation

The best case scenario is the slow convergence rate, while the worst case scenario is permanent entrapment to local minimum. An increased mutation rate which then slowly but steadily decreases leads to better exploration of the search space and steady transition to exploitation of the area of the global optimum. Also, the new termination criterion proved to be effective in avoiding useless estimations. In fact the decrease of the computational cost was 54.2% in the first example and 60.4%. Instead of using the best harmony found so far as it is, it is adjusted in the same way as the variables are adjusted in ALOPEX IV algorithm. With this adaptation we attain two things:

- Determinism in the adaptation of the best harmony, directing the best harmony to (possibly) new better one using vague knowledge of the objective function attained by the harmonies of the harmony memory. If the harmonies of the memory are concentrated at the area of the global optimum, then there is great possibility that the best harmony will be directed to it.

- The stochastic adjustment can help in avoiding the entrapment to local optima, which may come handy in the case of being driven deterministically.

5.1.3 Image Stitching

The Image Stitching method combines a new harmony search variant which is more advanced than the previous one with image downsizing in order to accelerate the

image registration process, which is crucial in image stitching. The method has the following advantages:

- Efficiency

- Automatic

- Simplicity

These factors make the method easy to use for doctors and biologist who are likely to have minimal expertise in software or image processing. The GUI, we presented above, is an example of a simple implementation of the use of powerful algorithms that can align images that do not have a grid positioning. Such a tool can be especially useful for non-experts, who seek to obtain high quality results.

5.1.4 Renyi divergence-based Mutual Information as an image comparison measure

The results of experiments, where mutual-information-like statistical measures based on Renyi divergence for image comparison were used, lead to the following conclusions: From Tables I-III, we conclude the following:

- The increase both of the α-parameter and the percentage reduces the mean absolute error (MAE), which means that the increase of these parameters makes the similarity measure more robust.

- The increase of the percentage of the pixel intensities increase on average the robustness of the similarity measure for any given value of the parameter α. This is intuitively expected, since the more information we have, the greater the possibility to find the optimum transformation. However, this increase is more pronounced in the higher values of α parameter and even more evident in the rotation accuracy of the rigid registration.

- The use of higher α parameter values increases the performance of the Renyi similarity measure. However, the increase is more pronounced in the smaller image sampling schemes and less in the larger.

From the numerical results, we observe that on average the increase of α reduces the MAE. Similarly, in Table V the increase of the sampling size leads to reduction of the MAE in rotation and translation.

The increased errors at lower values of parameter α can be attributed to fact that the ratios of the distributions are being equally weighed despite their dissimilarities (Eq. 3). The subsampling method can indeed reduce the duration of image registration at the expense of registration accuracy. In the case of medical image registration, where accuracy and speed (via subsampling) are potentially very important issues, the α parameter should be increased and in any case one should avoid smaller values of α as well as very small subsampling percentages. The choice of α around 1 seems to be optimal in most cases, however our work shows that in computationally demanding rigid registration tasks the studied parameters, if optimized, can lead to increased performance and registration accuracy

5.1.5 Surrogate Models based on Machine Learning

Surrogate models have never been used in metaheuristic-based image registration. This is an important step in reducing the computational cost in intensity-based image registration and bring them closer to feature-based methods in terms of duration. Specifically, the partial replacement of the original Mutual Information, (whose estimate for a given transformation is of the order of milliseconds which can be computationally expensive) by a computationally cheap and accurate approximation of it can decrease the duration of the image registration process up to 40% in the medical rigid image registration experiments. The mean duration of the estimation of the image similarity for a given rigid transformation is 47.55% despite the inclusion of the training process in the calculation. Because of the good results, it was decided to expand the experiments to other images of various sizes and various image data. The comparison of the approximation methods Kernel Recursive Least Squares, Kernel Ridge Regression, Multilayer Perceptron, Radial Basis Network and Epsilon-Insensitive Support Vector Regression has shown that Epsilon-Insensitive Support Vector Regression has the minimum average error. Also, the fact that the duration of each SVR estimation is of the order of $[10^{-6}, 10^{-5}]$ seconds, regardless the image size, is a good factor for the partial substitution of Mutual Information. The results of the image registration experiments have shown that Regular Step Gradient Descent and (1+1) Evolution strategy were not as successful as the Harmony Search variant. In fact, Regular Step Gradient Descent had the worst performance. My Harmony Search variant was successful with and without the use of SVR-based Surrogate Models. However, the use of Surrogate Models reduced the duration by 47%. This shows that Surrogate Models can be a new alternative strategy for the reduction of the computational cost, which can be used in conjunction with parallelism and subsampling, especially in more complex image transformations.

5.1.6 Affine Registration

The results of the experiments in affine image registration lead to the following conclusions:

- Elitist GA is the least successfull of the three metaheuristic optimization methods. The possible reasons for it are the following:

    - Blind Mutation: The mutation can be an important factor, since it facilitates the escape from local optima. However, mutation that is applied only for the sake of the mutation and without any information regarding the search space, may lead to escape from current local optima, only to be entrapped in new ones.

    - No proper exploitation of the area of the candidate solutions. Unlike the two Harmony Search variants, Elitist GA use blind mutation, which may help from escape from local optima, but it is absolutely blind.

- The first Harmony Search variant has better performance than the Elitist GA, but it is surpassed by the second Harmony Search variant. The reasons are the following ones:

    - If the new constructed solution is not better than the worst solution of the population, it is discarded as useless, even if it may have an element that can be used for the construction of a better solution.

    - The diversity of the construted solutions is less than that of the second variant.

- The second Harmony Search variant, both with and without the use of Support Vector Regression as a Surrogate Model, produced comparable results, although the use of the SVR Surrogate model shows a rapid decrease in computational cost up to 60%.

5.2   Open Issues for Future Research

Despite the new successful approaches regarding intensity-based image registration presented in this thesis, there are certain issues that need to be adressed:

1. Although the introduction of Surrogate models in intensity-based image registration has lead to a decrease in the computational cost, the need to use a significant part of the pixel values of the images makes them still slower than the feature-based ones.

2. Surrogate models, although successful, have only one, but major disadvantage: the computational cost of their construction. In the case of constructing a surrogate model of a function that has many arguments, a larger data sample for training is needed. This can lead to an increase in the computational cost in more complex image registration methods

3. Mutual information-based methods can be constructed via use of many different statistical divergences, whose perfomance depends on their respective parameters. As far as the literature is concerned, there is no research that shows which statistical divergence is optimal for all images regardless of their origins, lighting conditions and time that they are acquired.

Ergo, there is still room for research for the minimization of the computational cost of image registration process in intensity-based images to the level of the feature-based ones. The easiest approach that can be applied for further reduction is to combine:

- GPU Programming for the reduction of the computational cost of the original objective functon (e.g. Mutual Information) and the training of the surrogate model

- Subsampling

Also, another possibility is to fuse metaheuristics with conventional methods in order to create advanced metaheuristic methods that can easily locate global optima. In the case of image comparison, there is need for further research on the statistic divergences using a significantly large number of image pairs of various sources. Also, a new promised area is Deep Learning which has recently been used as a means for image comparison even more successful than intensity-based measures such as mutual information, although it can be used only on images similar to the training set. A grandiose challenge is to expand neural computation and design a new general neural method that can be used for every image pair regardless its origin and its nature. Being able to find a way to reduce of the computational cost of the intensity-based methods to the level of the feature-based ones can have a tremendous effect. The need for image preprocessing will be rendered obsolete. After all, image preprocessing may take some time to get good inter-mediate images for image registration and the preprocessing process may be applicable for certain image types only. Last but not least, it becomes even more difficult in the case of intermodal image registration. Without the need for preprocessing and with computational cost to the level of feature-based methods, intensity-based methods can be used for real time use such as medical surgery.

Bibliography

[1] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[2] Barbara Zitová and Jan Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

[3] Cr Maurer and Jm Fitzpatrick. A review of medical image registration. *Interactive image-guided . . .* , pages 1–49, 1993.

[4] Joseph Hajnal, David Hawkes, and Derek Hill. *Medical image registration*, volume 53. 2001.

[5] Denis P Shamonin, Esther E Bron, Boudewijn P F Lelieveldt, Marion Smits, Stefan Klein, and Marius Staring. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Frontiers in neuroinformatics*, 7(January):50, 2013.

[6] JB Antoine Maintz and Max A Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.

[7] Uwe Pietrzyk. Registration of mri and pet images for clinical applications. *Medical Image Registration*, 2001.

[8] Uwe Pietrzyk, Karl Herholz, Alexander Schuster, Hans-Martin v Stockhausen, Helmut Lucht, and Wolf-Dieter Heiss. Clinical applications of registration and fusion of multimodality brain images from pet, spect, ct, and mri. *European journal of radiology*, 21(3):174–182, 1996.

[9] Morten Bro-Nielsen. Medical Image Registration and Surgery Simulation. 1997.

[10] Joseph V Hajnal, DL Hill, and David J Hawkes. View of the future. *Medical Image Registration*, 2000.

[11] Daniel Rueckert and Paul Aljabar. Nonrigid registration of medical images: Theory, methods, and applications [applications corner]. *IEEE Signal Processing Magazine*, 27(4):113–119, 2010.

[12] Xueli Liu, Dongsheng Jiang, Manning Wang, and Zhijian Song. Image synthesis-based multi-modal image registration framework by using deep fully convolutional networks. *Medical & biological engineering &computing*, pages 1–12, 2018.

[13] Roger D Le Moigne, Jacqueline and Netanyahu, Nathan S and Eastman. *Image Registration for Remote Sensing*. Cambridge University Press, 2011.

[14] I De Falco, A Della Cioppa, D Maisto, U Scafuri, and E Tarantino. Satellite image registration by distributed differential evolution. In M Giacobini, editor, *Applications of Evolutionary Computing, Proceedings*, volume 4448, pages 251–260. Springer-Verlag Berlin, 2007.

[15] Changno Lee and James Bethel. Georegistration of airborne hyperspectral image data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7):1347–1351, 2001.

[16] Hector Erives and Glenn J Fitzgerald. Automatic Subpixel Registration for a Tunable Hyperspectral Imaging System. *IEEE Geoscience and remote sensing letters*, 3(3):397—-400, 2006.

[17] Christian Buil. Iris: Astronomical image-processing software. In *Digital Astrophotography: The State of the Art*, pages 79–88. Springer, 2005.

[18] John RG Townshend, Christopher O Justice, Charlotte Gurney, and James McManus. The impact of misregistration on change detection. *IEEE Transactions on Geoscience and remote sensing*, 30(5):1054–1060, 1992.

[19] Xiaolong Dai and Siamak Khorram. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote sensing*, 36(5):1566–1577, 1998.

[20] Richard Szeliski. Image Alignment and Stitching: A Tutorial. *Foundations and Trends® in Computer Graphics and Vision*, 2(1):1–104, 2006.

[21] Vladan Rankov, Rosalind J. Locke, Richard J. Edens, Paul R. Barber, and Borivoj Vojnovic. An Algorithm for image stitching and blending. In *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XII*, pages 190—-200, 2005.

[22] Mateusz Brzeszcz and Toby P. Breckon. Real-time construction and visualisation of drift-free video mosaics from unconstrained camera motion. *Journal of Engineering*, pages 1–12, 2015.

[23] Yuichi Motai and Akio Kosaka. Hand-Eye Calibration Applied to Viewpoint Selection for Robotic Vision. *IEEE Transactions on Industrial Electronics*, 55(10):3731–3741, 2008.

[24] Alvaro Collet, Dmitry Berenson, Siddhartha S Srinivasa, and Dave Ferguson. Object recognition and full pose registration from a single image for robotic manipulation. In *2009 IEEE International Conference on Robotics and Automation*, pages 48–55. IEEE, 2009.

[25] François Pomerleau, Francis Colas, Roland Siegwart, et al. A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends® in Robotics*, 4(1):1–104, 2015.

[26] Gaurav Pandey, James R McBride, Silvio Savarese, and Ryan M Eustice. Automatic extrinsic calibration of vision and lidar by maximizing mutual information. *Journal of Field Robotics*, 32(5):696–722, 2015.

[27] Fakhre Alam, Sami Ur Rahman, and Sehat Ullah. Medical image registration in image guided surgery: Issues, challenges and research opportunities. *Biocybernetics and Biomedical Engineering*, 38(1):71–89, 2017.

[28] Prachya Chalermwat. *High performance automatic image registration for remote sensing*. George Mason University Fairfax, VA, 1999.

[29] Junchen Wang, Hideyuki Suenaga, Kazuto Hoshi, Liangjing Yang, Etsuko Kobayashi, Ichiro Sakuma, and Hongen Liao. Augmented reality navigation with automatic marker-free image registration using 3-d image overlay for dental surgery. *IEEE transactions on biomedical engineering*, 61(4):1295–1304, 2014.

[30] J P W Pluim, J B a Maintz, and M a Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on medical imaging*, XX(Y):1–21, 2003.

[31] Guoli Song, Jianda Han, Yiwen Zhao, Zheng Wang, and Huibin Du. A Review on Medical Image Registration as an Optimization Problem. *Current Medical Imaging Reviews*, 13(3), 2017.

[32] W M Wells, P Viola, H Atsumi, S Nakajima, and R Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical image analysis*, 1(1):35–51, 1996.

[33] F Maes, A Collignon, D Vandermeulen, G Marchal, and P Suetens. Multimodality image registration by maximization of mutual information. *IEEE transactions on medical imaging*, 16:187–198, 1997.

[34] Darko Škerl, Boštjan Likar, and Franjo Pernuš. A protocol for evaluation of similarity measures for rigid registration. *IEEE Transactions on Medical Imaging*, 25(6):779–791, 2006.

[35] Hongxia Luan, Feihu Qi, Zhong Xue, Liya Chen, and Dinggang Shen. Multimodality image registration by maximization of quantitative-qualitative measure of mutual information. *Pattern Recognition*, 41(1):285–298, 2008.

[36] Daewon Lee, Matthias Hofmann, Florian Steinke, Yasemin Altun, Nathan D Cahill, and Bernhard Schölkopf. Learning Similarity Measure for Multi-Modal 3D Image Registration. In *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009*, pages 186—-193, 2009.

[37] Nathan D. Cahill. Normalized measures of mutual information with general definitions of entropy for multimodal image registration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6204 LNCS(8):258–268, 2010.

[38] M. Freiman, M. Werman, and L. Joskowicz. A curvelet-based patient-specific prior for accurate multi-modal brain image rigid registration. *Medical Image Analysis*, 15(1):125–132, 2011.

[39] Matthieu Ferrant, S Warfield, and Arya Nabavi. Registration of 3D intraoperative MR images of the brain using a finite element biomechanical model. *Medical Image Computing and Computer-Assited Intervention*, pages 249–258, 2000.

[40] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, 32(7), 2013.

[41] V R S Mani and S Arivazhagan. Survey of Medical Image Registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[42] 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.

[43] Zeinab Ghassabi, Jamshid Shanbehzadeh, Amin Sedaghat, and Emad Fatemizadeh. An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors. *EURASIP Journal on Image and Video Processing*, 2013(1), 2013.

[44] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF: Speeded Up Robust Features. In *European conference on computer vision*, pages 404–417, 2006.

[45] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. SURF : Speeded-Up Robust Features. *Computer vision and image understanding*, 110(3):346—-359, 2008.

[46] Xiaolong Dai and Siamak Khorram. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2351–2362, 1999.

[47] Mohamed Ali and David Clausi. Using the Canny edge detector for feature extraction and enhancement of remote sensing images. In *Geoscience and Remote Sensing Symposium, 2001. IGARSS'01. IEEE 2001 International*, pages 2298—-2300, 2001.

[48] Jinhui Tang, Xin Yang, Chunyan Liu, and Xiuqing Wu. Image Registration Based on Fitted Straight Lines of Edges. In *WCICA 2006. The Sixth World Congress on Intelligent Control and Automation, 2006*, pages 9782—-9785. IEEE, 2006.

[49] Chintan A Shah, Yongwei Sheng, and Laurence C Smith. Automated image registration based on pseudoinvariant metrics of dynamic land-surface features. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3908–3916, 2008.

[50] Anupama Gupta, • Harsh, K Verma, and Savita Gupta. A hybrid framework for registration of carotid ultrasound images combining iconic and geometric features. *Medical \& biological engineering \& computing*, 51(9):1043–1050, 2013.

[51] Maoguo Gong, Shengmeng Zhao, Licheng Jiao, Dayong Tian, and Shuang Wang. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7):4328–4338, 2014.

[52] Jyotirmoy Banerjee, Yuanyuan Sun, Camiel Klink, Renske Gahrmann, Wiro J Niessen, Adriaan Moelker, and Theo Van Walsum. Multiple-correlation similarity for block-matching based fast CT to ultrasound registration in liver interventions R. *Medical Image Analysis*, 53:132–141, 2019.

[53] 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.

[54] Philippe Thévenaz, Urs E Ruttimann, and Michael Unser. A Pyramid Approach to Subpixel Registration Based on Intensity. *IEEE transactions on image processing*, 7(1):27—-41, 1998.

[55] Frederik Maes, Dirk Vandermeulen, and Paul Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.

[56] George K Matsopoulos, Nicolaos A Mouravliansky, Konstantinos K Delibasis, and Konstantina S Nikita. Automatic Retinal Image Registration Scheme Using Global Optimization Techniques. *IEEE Transactions on Information Technology in Biomedicine*, 3(1):47—-60, 1999.

[57] Marcus Johansson. *Image Registration with Simulated Annealing and Genetic Algorithms*. PhD thesis, 2006.

[58] Andrea Valsecchi, Jérémie Dubois-Lacoste, Thomas Stützle, Sergio Damas, José Santamaria, and Linda Marrakchi-Kacem. Evolutionary medical image registration using automatic parameter tuning. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 1326–1333. IEEE, 2013.

[59] Constantinos Spanakis, Emmanuil Mathioudakis, Kampanis Nikos, Manolis Tsiknakis, and Kostas Marias. Elitism in intensity-based image registration. In *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 1–5, 2018.

[60] Nikolaus Hansen and Stefan Kern. Evaluating the cma evolution strategy on multimodal test functions. In *International Conference on Parallel Problem Solving from Nature*, pages 282–291. Springer, 2004.

[61] Susanne Winter, Bernhard Brendel, Ioannis Pechlivanis, Kirsten Schmieder, and Christian Igel. Registration of ct and intraoperative 3-d ultrasound images of the spine using evolutionary and gradient-based methods. *IEEE Transactions on Evolutionary Computation*, 12(3):284–296, 2008.

[62] 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.

[63] Taifeng Li, Quanke Pan, Gao Liang, and Peigen Li. Differential evolution algorithm-based range image registration for free-form surface parts quality inspection. *Swarm and Evolutionary Computation*, 36:106–123, 2017.

[64] Hadi Rezaei and Sassan Azadi. Nonrigid medical image registration using Hierarchical Particle Swarm Optimization. In *ICSCCW 2009. Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, 2009*, pages 1–4, 2009.

[65] Arpita Das and Mahua Bhattacharya. Affine-based registration of CT and MR modality images of human brain using multiresolution approaches: comparative study on genetic algorithm and particle swarm optimization. *Neural Computing and Applications*, 20(2):223–237, 2011.

[66] Yu Zhang and Honglei Zhou. Image stitching based on particle swarm and maximum mutual information algorithm. *Journal of Multimedia*, 8(5), 2013.

[67] Leonardo Rundo, Andrea Tangherloni, Carmelo Militello, Maria Carla Gilardi, and Giancarlo Mauri. Multimodal Medical Image Registration Using Particle Swarm Optimization. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–8, 2016.

[68] D R Sarvamangala and Raghavendra V Kulkarni. A Comparative Study of Bio-inspired Algorithms for Medical Image Registration. In *Advances in Intelligent Computing*, pages 27–44. 2019.

[69] 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.

[70] Jonathan Richard Shewchuk. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, 1994.

[71] Reeves Fletcher and Colin M Reeves. Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154, 1964.

[72] E Polak and G Ribiére. Note surla convergence de directions conjuges. *Rev. Francaise Informat Recherche Operationelle 3e Année*, 16(3):35–43, 1969.

[73] Magnus R Hestenes, Eduard Stiefel, et al. Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49(6):409–436, 1952.

[74] Yu-Hong Dai and Yaxiang Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10(1):177–182, 1999.

[75] Jack Kiefer. Sequential minimax search for a maximum. *Proceedings of the American mathematical society*, 4(3):502–506, 1953.

[76] Richard P. Brent. An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal*, 14(4):422–425, 1971.

[77] Fatiha Meskine, Miloud Chikr, El Mezouar, and Nasreddine Taleb. Nonsubsampled Contourlet Transform combined with Genetic Algorithms for Registration of Satellite Imaging. In *2009 3rd International Conference on Signals, Circuits and Systems (SCS)*, pages 1—-6. IEEE, 2009.

[78] Andrea Valsecchi, Sergio Damas, José Santamaría, and Linda Marrakchi-Kacem. Genetic Algorithms for Voxel-based Medical Image Registration. In *2013 IEEE Fourth International Workshop on Computational Intelligence in Medical Imaging (CIMI)*, pages 22—-29, 2013.

[79] Igor V Maslov and Izidor Gertner. Gradient-based genetic algorithms in image registration. In *Proc. SPIE*, page 510. 2001.

[80] H. Talbi, M. Batouche, and A. Draa. A quantum-inspired genetic algorithm for multi-source affine image registration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3211:147–154, 2004.

[81] Flavio Luiz Seixas, Luiz Satoru Ochi, Aura Conci, and Débora Muchaluat Saade. Image registration using genetic algorithms. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 1145—-1146, 2008.

[82] Y Fan and M Li. Optimization of projective transformation matrix in image stitching based on chaotic genetic algorithm. *International Journal of Intelligent Computing and Cybernetics*, 6(4):386–404, 2013.

[83] Martina Marinelli, Vincenzo Positano, Francesco Tucci, Danilo Neglia, and Luigi Landini. Automatic PET-CT Image Registration Method Based on Mutual Information and Genetic Algorithms. *The Scientific World Journal*, 2012:1–12, 2012.

[84] Jean-Michel Rouet, Jean-José Jacq, and Christian Roux. Genetic Algorithms for a Robust 3 - D MR - CT Registration. *IEEE transactions on information technology in biomedicine*, 4(2):126—-136, 2000.

[85] Xing Hua, Xiong * Zeng, Bo Qian, and Ren Xiang Wang. A remote sensing image subpixel matching combined genetic algorithm with least square matching. *Acta Geodaetica Et Cartographic Sinica*, 1, 2001.

[86] TS Douglas, SE Solomonidis, WA Sandham, and WD Spence. Ultrasound image matching using genetic algorithms. *Medical and Biological Engineering and Computing*, 40(2):168—-172, 2002.

[87] George C Kagadis, Konstantinos K Delibasis, George K Matsopoulos, Nikolaos A Mouravliansky, Pantelis A Asvestas, and George C Nikiforidis. A comparative study of surface-and volume-based techniques for the automatic registration between CT and SPECT brain images. *Medical Physics*, 29(2):201—-213, 2002.

[88] Renjie He and Ponnada a Narayana. Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 26(4):277–92, 2002.

[89] Chi Kin Chow, Hung Tat Tsui, and Tong Lee. Surface registration using a dynamic genetic algorithm. *Pattern recognition*, 37(1):105–117, 2004.

[90] Hongying Zhang, Xiaozhou Zhou, Jizhou Sun, and Jiawan Zhang. A Novel Medical Image Registration Method Based on Mutual Information and Genetic Algorithm. In *International Conference on Computer Graphics, Imaging and Vision: New Trends, 2005*, pages 221—-226, 2005.

[91] Famao Ye, Lin Su, and Shukai Li. Automatic multi-resolution image registration based on genetic algorithm and Hausdorff distance. *Chinese Optics Letters*, 4(7):386—-388, 2006.

[92] W.-Q. Li, C.-B. Wang, Q Wang, and G.-S. Chen. Tile image registration based on improved adaptive genetic algorithm. *Proceedings of the 2006 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV'06*, 1:216–222, 2006.

[93] Giuseppe Pascale and Luigi Troiano. A Niche Based Genetic Algorithm for Image Registration. In *ICEIS (2)*, pages 342—-347, 2007.

[94] Mahua Bhattacharya and Arpita Das. Multi resolution medical image registration using maximization of mutual information & optimization by genetic algorithm. In *IEEE Nuclear Science Symposium Conference Record, 2007. NSS'07.*, number September 2015, pages 2961–2964, 2007.

[95] Zsolt Jankó, Dmitry Chetverikov, and Anikó Ekárt. Using Genetic Algorithms in Computer Vision: Registering Images to 3D Surface Model *. *Acta Cybernetica*, 18(2):193–212, 2007.

[96] Yong Wang and Lu-ping Xu. A Global Optimized Registration Algorithm for Image Stitching. In *CISP'08. Congress on Image and Signal Processing, 2008*, volume 3, pages 525–529, 2008.

[97] Han-Ling Zhang and Fan Yang. Multimodality Medical Image Registration Using Hybrid Optimization Algorithm. In *BMEI 2008. International Conference on BioMedical Engineering and Informatics, 2008*, pages 183—-187, 2008.

[98] VT Ingole, CN Deshmukh, Anjali Joshi, and Deepak Shete. Medical image registration using genetic algorithm. In *2009 2nd International Conference on Emerging Trends in Engineering and Technology (ICETET)*, pages 63—-66, 2009.

[99] Xiaosheng Huang and Fang Zhang. Multi-modal Medical Image Registration Based on Gradient of Mutual Information and Hybrid Genetic Algorithm. In *2010 Third International Symposium on Intelligent Information Technology and Security Informatics (IITSI)*, pages 125—-128, 2010.

[100] Mahua Bhattacharya and Arpita Das. Multimodality medical image registration and fusion techniques using mutual information and genetic algorithm-based approaches. In *Software Tools and Algorithms for Biological Systems*, pages 441—-449. 2011.

[101] Hong Ying Qin. Research of Image Registration Based on Maximum Entropy Template Selection Algorithm. In *Advanced Materials Research*, pages 1138–1143, 2011.

[102] Andrea Valsecchi and Sergio Damas. An Image Registration Approach using Genetic Algorithms. pages 10–15, 2012.

[103] Andrea Valsecchi, Jérémie Dubois-Lacoste, Thomas Stützle, Sergio Damas, José Santamaría, and Linda Marrakchi-Kacem. Evolutionary Medical Image Registration using Automatic Parameter Tuning. In *2013 IEEE Congress on Evolutionary Computation (CEC)*, pages 1326–1333, 2013.

[104] S Shanmugapriya and S Poonguzhali. AN INTENSITY-BASED MEDICAL IMAGE REGISTRATION USING GENETIC ALGORITHM. *Signal \& Image Processing*, 5(3):53–58, 2014.

[105] Hasan Yetis, Mehmet Baygin, and Mehmet Karakose. A New Micro Genetic Algorithm Based Image Stitching Approach for Camera Arrays at Production Lines. In *5th International Conference on Manufacturing Engineering and Process (ICMEP), Istanbul, Turkey*, 2016.

[106] Hans-Paul Paul Schwefel. *Evolution and optimum seeking: the sixth generation*. John Wiley & Sons, Inc., 1993.

[107] Hans-Georg Beyer and Hans-Paul Schwefel. Evolution strategies – A comprehensive introduction. *Natural Computing*, 1(1):3 – 52, 2002.

[108] Stefan Klein. *Optimisation Methods for Medical Image Registration*. PhD thesis, Utrecht University, the Netherlands, 2008.

[109] Héctor Fernando Gómez García, Arturo González Vega, Arturo Hernández Aguirre, José Luis, Marroquín Zaleta, and Carlos Coello Coello. Robust Multiscale Affine 2D-Image Registration through Evolutionary Strategies. In *International Conference on Parallel Problem Solving from Nature*, pages 740–748. Springer, 2002.

[110] Susanne Winter, Bernhard Brendel, and Christian Igel. Registration of bone structures in 3D ultrasound and CT data: Comparison of different optimization strategies. In *International Congress Series*, volume 1281, pages 242–247. Elsevier, 2005.

[111] Stefan Klein, Marius Staring, and Josien P W Pluim. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. *IEEE T Image Processing*, 16(12):2879–90, 2007.

[112] Sean Gill, Parvin Mousavi, Gabor Fichtinger, Elvis Chen, Jonathan Boisvert, David Pichora, and Purang Abolmaesumi. Biomechanically constrained groupwise us to ct registration of the lumbar spine. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 803–810. Springer, 2009.

[113] Oscar Ibáñez, Lucia Ballerini, Oscar Cordón, Sergio Damas, and José Santamaría. An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification. *Information Sciences*, 179(23):3998–4028, 2009.

[114] Sean Gill, Parvin Mousavi, Gabor Fichtinger, David Pichora, and Purang Abolmaesumi. Group-wise registration of ultrasound to CT images of human vertebrae. In Michael I. Miga and Kenneth H. Wong, editors, *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, pages 1–9. International Society for Optics and Photonics, 2009.

[115] Siavash Khallaghi, Parvin Mousavi, Ren Hui Gong, Sean Gill, Jonathan Boisvert, Gabor Fichtinger, David Pichora, Dan Borschneck, and Purang Abolmaesumi. Registration of a Statistical Shape Model of the Lumbar Spine to 3D Ultrasound Images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 68–75. Springer, 2010.

[116] Dante De Nigris, D. Louis Collins, and Tal Arbel. Multi-modal image registration based on gradient orientations of minimal uncertainty. *IEEE Transactions on Medical Imaging*, 31(12):2343–2354, 2012.

[117] Yoshito Otake, Adam S Wang, J Webster Stayman, Ali Uneri, Gerhard Kleinszig, Sebastian Vogt, A Jay Khanna, Ziya L Gokaslan, and Jeffrey H Siewerdsen. Robust 3D-2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation. *Physics in Medicine & Biology*, 58(23):8535–8553, 2013.

[118] Y. Otake, R. J. Murphy, M. D. Kutzer, R. H. Taylor, and M. Armand. Piecewise-rigid 2D-3D registration for pose estimation of snake-like manipulator using an intraoperative x-ray projection. In Ziv R. Yaniv and David R. Holmes, editors, *Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9036, pages 1–6. International Society for Optics and Photonics, 2014.

[119] A Uneri, J W Stayman, T De Silva, A S Wang, G Kleinszig, S Vogt, A J Khanna, J.-P Wolinsky, Z L Gokaslan, and J H Siewerdsen. Known-Component 3D-2D Registration for Image Guidance and Quality Assurance in Spine Surgery Pedicle Screw Placement. In *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9415, pages 1–12, 2015.

[120] S Ouadah, J Webster Stayman, G Gang, Ali Uneri, T Ehtiati, and Jeffrey H Siewerdsen. Self-calibration of cone-beam ct geometry using 3d-2d image registration: Development and application to tasked-based imaging with a robotic c-arm. In *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9415, page 94151D. International Society for Optics and Photonics, 2015.

[121] S Ouadah, Joseph Webster Stayman, GJ Gang, T Ehtiati, and JH Siewerdsen. Self-calibration of cone-beam ct geometry using 3d–2d image registration. *Physics in Medicine & Biology*, 61(7):2613, 2016.

[122] M D Ketcha, T De Silva, A Uneri, G Kleinszig, S Vogt, J.-P Wolinsky, and J H Siewerdsen. Automatic Masking for Robust 3D-2D Image Registration in Image-Guided Spine Surgery. In Robert J. Webster III and Ziv R. Yaniv, editors, *Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling*. International Society for Optics and Photonics, 2016.

[123] T De Silva, A Uneri, M D Ketcha, S Reaungamornrat, G Kleinszig, S Vogt, N Aygun, S-F Lo, J-P Wolinsky, and J H Siewerdsen. 3D-2D image registration for target localization in spine surgery: investigation of similarity metrics providing robustness to content mismatch HHS Public Access. *Physics in Medicine & Biology*, 61(8):3009–3025, 2016.

[124] Catalina-Lucia Cocianu and Alexandru Stan. New Evolutionary-Based Techniques for Image Registration. *Applied Sciences*, 9(1):176, 2019.

[125] Rainer Storn and Kenneth Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.

[126] Michel Salomon, Guy-Ren Perrin, and Fabrice Heitz. Differential Evolution for Medical Image Registration. *International Conference on Artificial Intelligence*, pages 201–207, 2000.

[127] Hassiba Talbi and Mohamed Batouche. Hybrid particle swarm with differential evolution for multimodal image registration. In *Proceedings of the IEEE International Conference on Industrial Technology*, volume 3, pages 1567–1572, 2004.

[128] Xiaoyan Xu and Robert D Dony. Differential evolution with Powell's direction set method in medical image registration. In *2004 2nd IEEE International*

Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), pages 732–735. IEEE, 2004.

[129] Michel Salomon, Guy-René Perrin, Fabrice Heitz, and J-P Armspach. Parallel differential evolution: application to 3-d medical image registration. In *Differential evolution*, pages 353–411. 2005.

[130] Bartosz Telenczuk, J Ledesma-carbayo, Javier A Velazquez-muriel, Carlos O S Sorzano, Jose-maria Carazo, and Andres Santos. Molecular image registration using mutual information and differential evolution optimization. In *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006*, pages 844–847. IEEE, 2006.

[131] Zhenbang Hu, Wenyin Gong, and Zhihua Cai. Multi-resolution remote sensing image registration using differential evolution with adaptive strategy selection. *Optical Engineering*, 51(10):101707, 2012.

[132] José Santamaría, Sergio Damas, José M García-Torres, and Oscar Cordón. Self-adaptive evolutionary image registration using differential evolution and artificial immune systems q. *Pattern Recognition Letters*, 33(16):2065–2070, 2012.

[133] Mohsen Mirkhani, Rana Forsati, Alireza Mohammad Shahri, and Alireza Moayedikia. A novel efficient algorithm for mobile robot localization. *Robotics and Autonomous Systems*, 61(9):920–931, 2013.

[134] Ivanoe De Falco, Antonio Della Cioppa, Domenico Maisto, Umberto Scafuri, and Ernesto Tarantino. Adding chaos to differential evolution for range image registration. In *European Conference on the Applications of Evolutionary Computation*, number May 2014, pages 344–353. Springer, 2013.

[135] Wenping Ma, Xiafei Fan, Yue Wu, and Licheng Jiao. An orthogonal learning differential evolution algorithm for remote sensing image registration. *Mathematical Problems in Engineering*, 2014:16–22, 2014.

[136] Douglas C Montgomery. *Design and analysis of experiments*. John wiley & sons, 2017.

[137] Taifeng Li, Liang Gao, Quanke Pan, and Peigen Li. Differential Evolution Algorithm-based Range Image Registration with a Novel Point Descriptor. In *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, page 325, 2016.

[138] Yuan Qin, Haidong Hu, Yujiao Shi, Ye Liu, and Hao Gao. An Artificial Bee Colony Algorithm Hybrid with Differential Evolution for Multi-temporal Image Registration. In *2016 35th Chinese Control Conference (CCC)*, pages 2734–2739, 2016.

[139] Taifeng Li, Liang Gao, Quanke Pan, and Peigen Li. Differential evolution algorithm-based range image registration with scaling parameters. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 4508—-4512, 2016.

[140] Feiyi Xu, Haidong Hu, Hao Gao, Member Ieee, and Baoyun Wang. Multi-temporal Image Registration Utilizing a Differential Evolution Algorithm with Replacement Strategy. In *2016 Chinese Control and Decision Conference (CCDC)*, pages 752–757, 2016.

[141] Jason Brownlee. *Clever Algorithms: Nature-Inspired Programming Recipes*. Lulu, 2011.

[142] James Kennedy and Russell Eberhart. Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948. IEEE, 1995.

[143] Mark P. Wachowiak, Renata Smolíková, Yufeng Zheng, Jacek M. Zurada, and Adel S. Elmaghraby. An approach to multimodal biomedical image registration utilizing particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, 8(3):289–301, 2004.

[144] Isao Li, Qi and Sato. Multimodality Image Registration by Particle Swarm Optimization of Mutual Information. In *International Conference on Intelligent Computing*, pages 1120–1130, 2007.

[145] Jing Jin, Qiang Wang, and Yi Shen. High-performance medical image registration using improved particle swarm optimization. In *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, 2008.

[146] Yen-Wei Chen, Chen-Lun Lin, and Aya Mimori. Multimodal Medical Image Registration Using Particle Swarm Optimization. In *2008 Eighth International Conference on Intelligent Systems Design and Applications*, volume 1, pages 127–131, 2008.

[147] Wang Anna, Wang Tingjun, Zhang Jinjin, and Xue Silin. A Novel Method of Medical Image Registration Based on DTCWT and NPSO. *5th International Conference on Natural Computation, ICNC 2009*, 5(1):23–27, 2009.

[148] Nemir Ahmed Al-Azzawi, Harsa Amylia Mat Sakim, and Wan Ahmed K.Wan Abdullah. Fast free-form registration based on Kullback-Leibler distance for multimodal medical image. In *2010 IEEE Symposium on Industrial Electronics and Applications (ISIEA)*, pages 501–506, 2010.

[149] Di Zhou, Jun Sun, Choi-Hong Lai, Wenbo Xu, and Xiaoguang Lee. An improved quantum-behaved particle swarm optimization and its application to medical image registration. *International Journal of Computer Mathematics*, 88(6):1208—-1223, 2011.

[150] Fatemeh Ayatollahi, Shahriar Baradaran Shokouhi, and Ahmad Ayatollahi. A New Hybrid Particle Swarm Optimization for Multimodal Brain Image Registration. *Journal of Biomedical Science and Engineering*, 5(4):153–161, 2012.

[151] Chen-Lun Lin, Aya Mimori, and Yen-Wei Chen. Hybrid particle swarm optimization and its application to multimodal 3D medical image registration. *Computational intelligence and neuroscience*, 2012:561406, 2012.

[152] Fang Liu, Haibin Duan, and Yimin Deng. A Chaotic Quantum-behaved Particle Swarm Optimization based on Lateral Inhibition for Image Matching. *Optik-International Journal for Light and Electron Optics*, 123(21):1955–1960, nov 2012.

[153] Leandro dos Santos Coelho. A quantum particle swarm optimizer with chaotic mutation operator. *Chaos, Solitons & Fractals*, 37(5):1409–1418, 2008.

[154] Lydia Schwab, Manuel Schmitt, and Rolf Wanka. Multimodal medical image registration using particle swarm optimization with influence of the data's initial orientation. In *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1—-8, 2015.

[155] Smita Pradhan and Dipti Patra. RMI based non-rigid image registration using BF-QPSO optimization and P-spline. *AEU - International Journal of Electronics and Communications*, 69(3):609–621, 2015.

[156] Youwen Zhuang, Kun Gao, Xianghu Miu, Lu Han, and Xuemei Gong. Infrared and visual image registration based on mutual information with a combined particle swarm optimization - Powell search algorithm. *Optik - International Journal for Light and Electron Optics*, 127(1):188–191, 2016.

[157] DR Sarvamangala and Raghavendra V Kulkarni. Swarm Intelligence Algorithms for Medical Image Registration: A Comparative Study. In *International Conference on Computational Intelligence, Communications, and Business Analytics*, pages 451–465. Springer, 2017.

[158] Mohamed Abdel-Basset, Ahmed E. Fakhry, Ibrahim El-henawy, Tie Qiu, and Arun Kumar Sangaiah. Feature and Intensity Based Medical Image Registration Using Particle Swarm Optimization. *Journal of Medical Systems*, 41(12), 2017.

[159] Chengjia Wang, Keith A. Goatman, James Boardman, Erin Beveridge, David Newby, and Scott Semple. Distance Oriented Particle Swarm Optimizer for Brain Image Registration. *IEEE Access*, 7:56016–56027, 2019.

[160] Yudong Zhang and Lenan Wu. A novel method for rigid image registration based on firefly algorithm. *International Journal of Research and Reviews in Soft and Intelligent Computing (IJRRSIC)*, 2(2), 2012.

[161] Enrique Bermejo, Andrea Valsecchi, Sergio Damas, and Oscar Cord. Bacterial Foraging Optimization For Intensity-based Medical Image Registration. In *2015 IEEE Congress on Evolutionary Computation (CEC)*, pages 2436–2443. IEEE, 2015.

[162] Enrique Bermejo, Oscar Cordón, Sergio Damas, and Jose Santamar\'\ia. A comparative study on the application of advanced bacterial foraging models to image registration. *Information Sciences*, 295:160–181, 2015.

[163] Christoph Vetter, Christoph Guetter, Chenyang Xu, and Rüdiger Westermann. Non-rigid multi-modal registration on the GPU. *Proceedings of SPIE*, 6512:651228–651228–8, 2007.

[164] Daniel Henrik Adler. Accelerated Medical Image Registration using the Graphics Processing Unit. 2011.

[165] Yiqi Cai, Xiaohu Guo, Zichun Zhong, and Weihua Mao. Dynamic meshing for deformable image registration. *CAD Computer Aided Design*, 58, 2015.

[166] Antonio Ruiz, Antonio Ruiz, Manuel Ujaldon, Manuel Ujaldon, Lee Cooper, Lee Cooper, Kun Huang, and Kun Huang. Non-rigid Registration for Large Sets of Microscopic Images on Graphics Processors. *Journal of Signal Processing Systems*, 55(1-3):229–250, 2008.

[167] William Plishker, Omkar Dandekar, Shuvra S Bhattacharyya, and Raj Shekhar. Utilizing Hierarchical Multiprocessing for Medical Image Registration. *Journal of Chemical Information and Modeling*, 27(2):61—-68, 2010.

[168] Yixun Liu, Andriy Kot, Fotis Drakopoulos, Chengjun Yao, Andriy Fedorov, Andinet Enquobahrie, Olivier Clatz, and Nikos P. Chrisochoides. An ITK implementation of a physics-based non-rigid registration method for brain deformation in image-guided neurosurgery. *Frontiers in Neuroinformatics*, 8:33, 2014.

[169] Prachya Chalermwat, Tarek El-Ghazawi, and Jacqueline Lemoigne. GA-based parallel image registration on parallel clusters. *Parallel and Distributed Processing*, pages 257—-265, 1999.

[170] Prachya Chalermwat, Tarek El-Ghazawi, and Jacqueline Lemoigne. 2-phase GA-based image registration on parallel clusters. *Future Generation Computer Systems*, 17(4):467–476, 2001.

[171] Torsten Butz and Jean-Philippe Thiran. Affine Registration with Feature Space Mutual Information. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 549—-556, 2001.

[172] Fumihiko Ino, Kanrou Ooyama, and Kenichi Hagihara. A data distributed parallel algorithm for nonrigid image registration. *Parallel Computing*, 31(1):19–43, 2005.

[173] M. P Wachowiak and T. M Peters. High-performance medical image registration using new optimization techniques. *IEEE Transactions on Information Technology in Biomedicine*, 10(2):344–353, 2006.

[174] I. De Falco, D. Maisto, U. Scafuri, E. Tarantino, and A. Della Cioppa. Distributed differential evolution for the registration of remotely sensed images. In *Proceedings - 15th EUROMICRO International Conference on Parallel, Distributed and Network-Based Processing, PDP 2007*, pages 358–362, 2007.

[175] Mohammed Yagouni. Using of Metaheuristics for Optimizing Satellite Image Registration. In *CIE 2009. International Conference on Computers & Industrial Engineering, 2009*, pages 531–535. IEEE, 2009.

[176] Torsten Butz and Jean Philippe Thiran. Affine registration with feature space mutual information. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 549–556, 2001.

[177] Wen Peng, Ruofeng Tong, Guiping Qian, and Jinxiang Dong. *A Constrained Ant Colony Algorithm for Image Registration*, pages 1–11. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[178] Silviu Ioan Bejinariu, Hariton Costin, Florin Rotaru, Ramona Luca, and Cristina Nit. Parallel Processing and Bio-inspired Computing for Biomedical Image Registration. 22(2):253–278, 2014.

[179] Fotis Drakopoulos and Nikos P. Chrisochoides. Accurate and fast deformable medical image registration for brain tumor resection using image-guided neurosurgery. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 4(2):112–126, 2016.

[180] Edward Castillo. Quadratic Penalty Method for Intensity-Based Deformable Image Registration and 4DCT Lung Motion Recovery. *Medical physics*, 2019.

[181] Joseph V Hajnal and Derek LG Hill. Medical image registration. *Physics in Medicine and Biology*, 46:1–45, 2001.

[182] Gary E Christensen, Richard D Rabbitt, Michael I Miller, and Others. Deformable templates using large deformation kinematics. *IEEE transactions on image processing*, 5(10):1435–1447, 1996.

[183] Jean-Philippe Thirion. Non-Rigid Matching Using Demons. In *Proceedings CVPR'96, 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1996.*, pages 245–251, 1996.

[184] Siavash Zokai and George Wolberg. Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations. *IEEE Transactions on Image Processing*, 14(10):1422–1434, 2005.

[185] D Knaan and L Joskowicz. LNCS 2878 - Effective Intensity-Based 2D/3D Rigid Registration between Fluoroscopic X-Ray and CT. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pages 351—-358, 2003.

[186] L{\'a}szl{\'o} G Ny{\'u}l, Jayaram K Udupa, and Punam K Saha. Incorporating a measure of local scale in voxel-based 3-D image registration. *IEEE Transactions on Medical Imaging*, 22(2):228–237, 2003.

[187] Alexis Roche, Gr̃goire Malandain, Xavier Pennec, and Nicholas Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1115–1124. Springer, 1998.

[188] J. P W Pluim, J. B Antoine Maintz, and M. a. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19(8):809–814, 2000.

[189] Frederik Maes, Dirk Vandermeulen, and Paul Suetens. Medical image registration using mutual information. *Proceedings of the IEEE*, 91(10):1699–1721, 2003.

[190] Andrea Valsecchi-European and Sergio Damas. Evolutionary intensity-based medical image registration: a review. *Current Medical Imaging Reviews*, 9(4):283—-297, 2013.

[191] Mert R Sabuncu and Peter J Ramadge. Gradient based optimization of an EMST image registration function. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, number 1, pages ii—-253, 2005.

[192] Andrea Valsecchi, Sergio Damas, José Santamaría, and Linda Marrakchi-Kacem. Intensity-based image registration using scatter search. *Artificial Intelligence in Medicine*, 60(3):151–163, 2014.

[193] Nicola Ritter, Robyn Owens, James Cooper, Robert H Eikelboom, and Paul P Van Saarloos. Registration of Stereo and Temporal Images of the Retina. *IEEE Transactions on medical imaging*, 18(5):404—-418, 1999.

[194] N M Alpert, D Berdichevsky, Z Levin, E D Morris, and A J Fischman. Improved Methods for Image Registration. *NeuroImage*, 3(1):10–18, 1996.

[195] Martin Styner. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Transactions on Medical Imaging*, 19(3):153–165, 2000.

[196] Joaquim Salvi, Carles Matabosch, David Fofi, and Josep Forest. A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, 25(5):578–596, 2007.

[197] Johanöfverstedt Johan" Johanöfverstedt, Joakim Lindblad, and Nataša Sladoje. Fast and Robust Symmetric Image Registration Based on Intensity and Spatial Information. *arXiv preprint arXiv:1807.11599*, 2018.

[198] Josien P W Pluim, J B Antoine Maintz, and Max A Viergever. Mutual information matching and interpolation artefacts. In *Medical Imaging 1999: Image Processing*, volume 3661, pages 56–66, 1999.

[199] Josien P. W. Pluim, J. B. Antoine Maintz, and Max a. Viergever. Interpolation Artefacts in Mutual Information-Based Image Registration. *Computer Vision and Image Understanding*, 77:211–232, 2000.

[200] Stefan Klein, Marius Staring, Keelin Murphy, Max A Viergever, and Josien P W Pluim. Elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.

[201] Maumita Bhattacharya. Evolutionary Approaches to Expensive Optimisation. *IJARAI) International Journal of Advanced Research in Artificial Intelligence*, 2(3):53–59, 2013.

[202] Alexander E I Brownlee, John R Woodward, and Jerry Swan. Metaheuristic Design Pattern: Surrogate Fitness Functions. In *Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 1261–1264. ACM, 2015.

[203] Mark P Wachowiak, Renata Smolíková, Georgia D Tourassi, and Adel S Elmaghraby. Similarity metrics based on nonadditive entropies for 2D-3D multimodal biomedical image registration. In *Medical Imaging 2003: Image Processing*, pages 1090–1101, 2003.

[204] J P Pluim, J B Maintz, and M A Viergever. F-information measures in medical image registration. *IEEE Trans Med Imaging*, 23(12):1508–1516, 2004.

[205] Hamid Krim, Yun He, and A Ben Hamza. A generalized divergence measure for robust image registration. *IEEE Transactions on Signal Processing*, 51(5):1211–1220, 2016.

[206] Ming-Chang Chiang, Rebecca A Dutton, Kiralee M Hayashi, Arthur W Toga, Oscar L Lopez, Howard J Aizenstein, James T Becker, and Paul M Thompson. Fluid registration of medical images using Jensen-Renyi divergence reveals 3D profile of brain atrophy in HIV/AIDS. In *3rd IEEE International Symposium on Biomedical Imaging Macro to Nano 2006*, pages 193–196, 2006.

[207] Fei Wang, Tanveer Syeda-Mahmood, Baba C Vemuri, David Beymer, and Anand Rangarajan. Closed-form Jensen-Renyi divergence for mixture of Gaussians and applications to group-wise shape registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 648–655, 2009.

[208] András P. Keszei, Benjamin Berkels, and Thomas M. Deserno. Survey of Non-Rigid Registration Tools in Medicine, 2017.

[209] Zong Woo Geem, Joong Hoon Kim, and G.V Loganathan. A New Heuristic Optimization Algorithm: Harmony Search. *Simulation*, 76(2):60–68, 2001.

[210] Xin-She Yang. Harmony Search as a Metaheuristic Algorithm. In *Music-inspired harmony search algorithm*, volume 191, pages 1–14. Springer, 2009.

[211] Constantinos Spanakis, Emmanuil Mathioudakis, Manolis Tsiknakis, Kostas Marias, et al. Elitism in intensity-based image registration. In *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 1–5. IEEE, 2018.

[212] Constantinos Spanakis, Emmanuel Mathioudakis, Nikos Kampanis, Manolis Tsiknakis, and Kostas Marias. A new approach in robust image registration. In *2016 IEEE International Conference on Imaging Systems and Techniques (IST)*, number October, pages 449—-453, 2016.

[213] Constantinos Spanakis, Manolis Mathioudakis, Emmanuel Tsiknakis, Nikos Kampanis, and Kostas Marias. Function Approximation for Medical Image Registration. In *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, pages 1–5. IEEE, 2018.

[214] Ben Appleton, Andrew P Bradley, and Michael Wildermoth. Towards Optimal Image Stitching for Virtual Microscopy. In *Digital Image Computing: Techniques and Applications, 2005. DICTA'05. Proceedings 2005*, pages 44–51. IEEE, 2005.

[215] Dirk Steckhan, Tobias Bergen, Thomas Wittenberg, and Stephan Rupp. Efficient large scale Image Stitching for Virtual Microscopy. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4019–4023. IEEE, 2008.

[216] Dirk Schlierkamp-Voosen and H Mühlenbein. Predictive models for the breeder genetic algorithm. *Evolutionary Computation*, 1(1):25–49, 1993.

[217] Brad L M Ille, David E Goldberg, Brad L Miller, and David E Goldb. Genetic Algorithms , Tournament Selection, and the Effects of Noise. *Complex systems*, 9(3):193–212, 1995.

[218] Jinghui Zhong, Xiaomin Hu, Jun Zhang, and Min Gu. Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms. In *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, volume 2, pages 1115–1121, 2005.

[219] Tobias Blickle and Lothar Thiele. A Comparison of Selection Schemes used in Genetic Algorithms. *Evolutionary Computation*, 2(11):311–347, 1995.

[220] Zsolt Jankó, Dmitry Chetverikov, and Anikó Ekárt. Using a genetic algorithm to register an uncalibrated image pair to a 3d surface model. *Engineering Applications of Artificial Intelligence*, 19(3):269–276, 2006.

[221] Luciano Silva, Olga Regina Pereira Bellon, and Kim 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.

[222] Rakesh Singhai and Jyoti Singhai. Registration of Satellite Imagery Using Genetic Algorithm. II:2–7, 2012.

[223] Chi Kin Chow, Hung Tat Tsui, and Tong Lee. Surface registration using a dynamic genetic algorithm. *Pattern Recognition*, 37(1):105–117, 2004.

[224] Evgeny Lomonosov, Dmitry Chetverikov, and Anikó Ekárt. Pre-registration of arbitrarily oriented 3d surfaces using a genetic algorithm. *Pattern Recognition Letters*, 27(11):1201–1208, 2006.

[225] Craig Robertson and Robert B Fisher. Parallel evolutionary registration of range data. *Computer Vision and Image Understanding*, 87(1-3):39–50, 2002.

[226] S. Damas, O. Cordón, and J. Santamara. Medical image registration using evolutionary computation: An experimental survey. *IEEE Computational Intelligence Magazine*, 6(4):26–42, 2011.

[227] Keshvad Shahrivar. *New Developments in Evolutionary Image Registration for Complex 3D Scenarios*. PhD thesis, Universidad de Granada, 2017.

[228] J. Melorose, R. Perroy, and S. Careas. *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, volume 1. 2015.

[229] L Shi and K Rasheed. A survey of fitness approximation methods applied in evolutionary algorithms. In *Computational intelligence in expensive optimization problems*, pages 3–28. Springer, 2010.

[230] Justin Buhendwa Nyenyezi and Annick Sartenaer. Benchmarking some iterative linear systems solvers for deformable 3D images registration. *Welcome to ORBEL 32!*, page 71, 2018.

[231] 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.

[232] O Cordon, S Damas, and J Santamaria. A CHC evolutionary algorithm for 3D image registration. *Fuzzy Sets and Systems - Ifsa 2003, Proceedings*, 2715:404–411, 2003.

[233] G Troglio, J A Benediktsson, S B Serpico, G Moser, R A Karlsson, G H Halldorsson, and E Stefansson. Automatic registration of retina images based on genetic techniques. In *EMBS 2008. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008*, pages 5419—-5424, 2008.

[234] Zsolt Jank{\'o}, Dmitry Chetverikov, and Anik{\'o} Ek{\'a}rt. Using a genetic algorithm to register an uncalibrated image pair to a 3D surface model. *Engineering Applications of Artificial Intelligence*, 19(3):269—-276, 2006.

[235] Mahamed G H Omran and Mehrdad Mahdavi. Global-best harmony search. *Applied Mathematics and Computation*, 198(2):643–656, 2008.

[236] Yaochu Jin and Bernhard Sendhoff. Fitness approximation in evolutionary computation-A survey. In *Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation*, pages 1105–1112. Morgan Kaufmann Publishers Inc., 2002.

[237] Yaochu Jin, Senior Member, Markus Olhofer, and Bernhard Sendhoff. A Framework for Evolutionary Optimization With Approximate Fitness Functions. *IEEE Transactions on evolutionary computation*, 6(5):481–494, 2002.

[238] Y Jin. A comprehensive survey of fitness approximation in evolutionary computation. *Soft Computing Journal*, 9(1):3–12, 2005.

[239] Hongguang Li, Wenrui Ding, Xianbin Cao, and Chunlei Liu. Image registration and fusion of visible and infrared integrated camera for medium-altitude unmanned aerial vehicle remote sensing. *Remote Sensing*, 9(5):1–29, 2017.

[240] Daniel S Gareau, Yongbiao Li, Billy Huang, Zach Eastman, and Milind Rajadhyaksha. Confocal Mosaicing Microscopy in Mohs skin excisions: feasibility of rapid surgical pathology NIH Public Access. *Journal of Biomedical Optics*, 13(5):54001, 2008.

[241] Peter A Santi, Shane B Johnson, Matthias Hillenbrand, Patrick Z GrandPre, Tiffany J Glass, and James R Leger. Thin-sheet laser Imaging Microscopy for Optical Sectioning of Thick Tissues. *Biotechniques*, 46(4):287–294, 2009.

[242] Stephan Preibisch, Stephan Saalfeld, and Pavel Tomancak. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics*, 25(11):1463–1465, 2009.

[243] Erwin Schindelin, Johannes Arganda-Carreras, Ignacio Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Curtis Preibisch, Stephan Rueden, Stephan Saalfeld, Benjamin Schmid, and Others. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676, 2012.

[244] Terry S Yoo, Michael J Ackerman, William E Lorensen, Will Schroeder, Vikram Chalana, Stephen Aylward, Dimitris Metaxas, and Ross Whitaker. Engineering and algorithm design for an image processing API: a technical report on ITK-the insight toolkit. *Studies in health technology and informatics*, pages 586–592, 2002.

[245] Oleg Ponomarenko, Nikolay Ieremeiev, Vladimir Lukin, Lina Jin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, and Federica Battisti. A new color image database TID2013: Innovations and results. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 402–413, 2013.

[246] E Harth et al. Alopex: A stochastic method for determining visual receptive fields. 1974.

[247] Evangelia Tzanakou, R Michalak, and Erich Harth. The alopex process: visual receptive fields by response feedback. *Biological Cybernetics*, 35(3):161–174, 1979.

[248] M S Zakynthinaki and Yannis G Saridakis. Stochastic optimization for adaptive real-time wavefront correction. *Numerical Algorithms*, 33(1-4):509–520, 2003.

[249] MS Zakynthinaki, RO Barakat, CA Cordente Martínez, and J Sampedro Molinuevo. Stochastic optimization for the detection of changes in maternal heart rate kinetics during pregnancy. *Computer Physics Communications*, 182(3):683–691, 2011.

[250] Evangelia Miche Tzanakou. *Supervised and unsupervised pattern recognition: feature extraction and computational intelligence*. CRC press, 1999.

[251] Abhijit S Pandya, Ercan Sen, and Sam Hsu. Buffer allocation optimization in atm switching networks using alopex algorithm. *Neurocomputing*, 24(1-3):1–11, 1999.

[252] Paris N Stratis, George P Karatzas, Elena P Papadopoulou, Maria S Zakynthinaki, and Yiannis G Saridakis. Stochastic optimization for an analytical model of saltwater intrusion in coastal aquifers. *PloS one*, 11(9):e0162783, 2016.

[253] Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

[254] Sven F Crone ', Jose Guajardo^, and Richard Weber^. A study on the ability of Support Vector Regression and Neural Networks to Forecast Basic Time Series Patterns. In *IFIP International Conference on Artificial Intelligence in Theory and Practice*, pages 149—-158, 2006.

[255] Debasish Basak, Srimanta Pal, and Dipak Chandra Patranabis. Support Vector Regression. *Neural Information Processing – Letters and Reviews*, 11(10):203—-224, 2007.

[256] Roman M Balabin and Ekaterina I Lomakina. Support vector machine regression (SVR/LS-SVM)—an alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. *Analyst*, 136(8):1703–1712, 2011.

[257] Chen-Rui Chou. *Regression learning for 2D/3D image registration*. PhD thesis, The University of North Carolina at Chapel Hill, 2013.

[258] Yaakov Engel, Shie Mannor, and Ron Meir. The kernel recursive least-squares algorithm. *IEEE Transactions on signal processing*, 52(8):2275–2285, 2004.

[259] Ryan M Rifkin, Ross A Lippert, and D E Shaw Research. Notes on Regularized Least Squares. Technical report, 2007.

[260] Vladimir Vovk. Kernel ridge regression. In *Empirical inference*, pages 105—-116. 2013.

[261] David S Broomhead and David Lowe. Radial basis functions, multi-variable functional interpolation and adaptive networks. Technical report, Royal Signals and Radar Establishment Malvern (United Kingdom), 1988.

[262] Vladimir Vapnik. *Statistical Learning Theory*. 1998.

[263] http://www.civm.duhs.duke.edu/pubs/supplemental/NeuroImage200702/index.html.

[264] https://remotesensing.usgs.gov/gallery/.