UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Optimal transport applied to eye fundus image registration

Danilo Andrade Motta

Tese de Doutorado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Danilo Andrade Motta

Optimal transport applied to eye fundus image registration

Doctoral dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Doctorate Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Afonso Paiva Neto

USP – São Carlos
January 2019

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

M921o Motta, Danilo
Optimal transport applied to eye fundus image
registration / Danilo Motta; orientador Afonso
Paiva. -- São Carlos, 2019.
92 p.

Tese (Doutorado - Programa de Pós-Graduação em
Ciências de Computação e Matemática Computacional) --
Instituto de Ciências Matemáticas e de Computação,
Universidade de São Paulo, 2019.

1. Optimal transport. 2. Image processing. 3.
Graph matching. 4. Registration. 5. Medical images.
I. Paiva, Afonso, orient. II. Título.

Danilo Andrade Motta

**Transporte ótimo de massa aplicado ao registro de imagens
de fundo do olho**

Tese apresentada ao Instituto de Ciências
Matemáticas e de Computação – ICMC-USP,
como parte dos requisitos para obtenção do título
de Doutor em Ciências – Ciências de Computação e
Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e
Matemática Computacional

Orientador: Prof. Dr. Afonso Paiva Neto

USP – São Carlos
Janeiro de 2019

To Carina, always my support and inspiration.

ACKNOWLEDGEMENTS

I want to thank my colleagues who have ever supported me and collaborated with me to make this research possible.

To my advisor Afonso Paiva who helped me to take my first steps into research and who have always encouraged my quest to find new interesting problems to work on, that have ever made time to talk with me and widen my views. Afonso always treated me as a colleague and a friend, and I could never have asked for more.

I also want to thank Carlos Scheiddeger who put his trust in my work, and that I'm forever grateful for his support during my time in Tucson. From all the things Carlos taught me, I think the most important one was to realize how amazing it is what we do.

To Paulo Pagliosa my undergraduate advisor who has started all of this.

To Wallace Casaca and Erick Gomez-Nieto, I'm grateful that I made such wonderful friends be able to have fun while pulling all-nighters before submission deadlines.

I also want to thank Gustavo Nonato, Matthew Berger, and Josh Levine for the opportunity of working with such amazing researchers during my Ph.D. who taught me that we must seek for the best but also that science is made more of questions than results.

To all the teachers and employees from the ICMC-USP that patiently answered my many questions over the years.

I also want to acknowledge the funding sources for my research during these years, including the support of the agencies CAPES and CNPq.

To my family and friends who made the weight lighter all the way.

*“The most exciting phrase to hear in science,
the one that heralds new discoveries,
is not, ‘Eureka! I’ve found it,’ but,
‘That’s funny!’”*
(Isaac Asimov)

RESUMO

MOTTA, DANILO. **Transporte ótimo de massa aplicado ao registro de imagens de fundo do olho**. 2019. 92 p. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

O transporte ótimo se tornou uma ferramenta promissora e eficaz para apoiar o processamento de imagens moderno, processamento geométrico e até aprendizado de máquina. De fato, a teoria do transporte ótimo permite uma grande flexibilidade na modelagem de problemas, pois diferentes recursos de otimização podem ser empregados enquanto se preserva uma propriedade relevante ao contexto que pode ser interpretada como massa. Nesta pesquisa, nós introduzimos uma nova técnica automática para o registro da imagem do fundo do olho que é baseada na teoria ótima do transporte, filtros de processamento de imagem, correspondência de grafos e transformações geométricas em uma estrutura concisa e unificada. Dadas duas imagens de fundo ocular, construímos grafos representativos que incorporam em suas estruturas informações espaciais e topológicas dos vasos sanguíneos do olho. Os grafos produzidos são usados como entrada pelos nossos modelo de transporte ótimo, a fim de estabelecer uma correspondência entre seus conjuntos de nós. Propomos também uma nova medida que estima a qualidade do registro e uma extensão de uma tecnica de remoção de outliers chamada DeSAC. Finalmente, transformações geométricas são realizadas entre as imagens para realizar adequadamente a tarefa de registro. Nosso método baseia-se em uma sólida base matemática, é fácil de implementar e funciona bem lidando com outliers criados durante o estágio de correspondência, produzindo soluções determinísticas e precisas. Demonstramos a exatidão e eficácia da metodologia proposta por meio de uma abordagem abrangente de comparações qualitativas e quantitativas contra vários métodos representativos do estado da arte em diferentes bases de dados de imagens de fundo de olho.

Palavras-chave: Transporte ótimo, processamento de imagem, correspondência de grafos, registro, imagens médicas.

ABSTRACT

MOTTA, DANILO. **Optimal transport applied to eye fundus image registration**. 2019. 92 p. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Optimal transport has emerged as a promising and effective tool for supporting modern image processing, geometric processing, and even machine learning. Indeed, the optimal transport theory enables great flexibility in modeling problems, as different optimization resources can be successfully employed while preserving a context relevant property that can be interpreted as mass. In this research, we introduce a novel automatic technique for eye fundus image registration which is based on optimal transport theory, image processing filters, graph matching, and geometric transformations into a concise and unified framework. Given two ocular fundus images, we construct representative graphs which embed in their structures spatial and topological information from the eye's blood vessels. The graphs produced are then used as input by our optimal transport model in order to establish a correspondence between their sets of nodes. We also proposed a new measure that estimates the register quality and an extension of an outlier removal technique called DeSAC. Finally, the best geometric transformation is performed on the image to properly accomplish the registration task. Our method relies on a solid mathematical foundation, is easy-to-implement and performs well when dealing with outliers created during the matching stage, producing deterministic and accurate solutions. We demonstrate the accuracy and effectiveness of the proposed methodology through a comprehensive set of qualitative and quantitative comparisons against various representative state-of-the-art methods on different fundus image databases.

Keywords: Optimal transport, image processing, graph matching, registration, medical images.

LIST OF FIGURES

Figure 1 – Examples of eye fundus images.	21
Figure 2 – Examples of eye fundus images pairs, as columns, with noise from patients with disease-related deformations and illumination artifacts.	23
Figure 3 – 2D Optimal transport problem.	24
Figure 4 – Illustration of the optimal transport problem.	28
Figure 5 – Illustration of some distinct distributions that can be used in an optimal transport problem: discrete distributions where $\nu = a_i \delta_{x_i}$, from which uniform are blue and $a_i = \frac{1}{n}$ and arbitrary are in green where the area/length can be interpreted as a_i ; in red can be seen continuous distributions where $d\nu(x) = \rho(x)dx$	30
Figure 6 – Pipeline overview of VOTUS.	39
Figure 7 – Graph construction (from left to right): blood vessel binarization, blood vessel skeletonization, full graph after skeletization and its simplification drawn on the original images.	40
Figure 8 – Optimal graph matching: the matrix $\mathbf{X}$ encodes the assignment between the nodes $\mathcal{N}_I$ (rows/orange) and $\mathcal{N}_R$ (columns/green), including their associated ghost nodes $\mathbf{g}_I$ (light orange) and $\mathbf{g}_R$ (light green).	43
Figure 9 – Neighbor distance between the nodes $\mathbf{v}_I$ and $\mathbf{v}_R$	44
Figure 10 – Example of block features extracted in two pair images ordered in columns: first are shown the keypoints origin; the gray-image patch extracted is in the second column, and gabor filter patches in the last column.	46
Figure 11 – Saliency-based vector (from left to right): input image, the most salient region (gray), its center point (in black), and vector field (yellow) built in terms of the ONH center point and the graph nodes.	47
Figure 12 – Outlier filtering: before filtering (yellow) with outliers (cyan) and after the improved DeSAC filtering (white).	50
Figure 13 – Illustrative example of the different characteristics of the transformations considered.	53
Figure 14 – Measuring the registration accuracy from the GC metric. First, blood vessels B_R and B_M are built from the fundus images I_R and I_M . Then, a geometric transformation model T is applied on B_M so as to generate $T(B_M)$. Finally, the Gain Coefficient is computed as in Equation 5.4.	55

Figure 15 – Illustration of our GC-based approach for selecting the optimal transformation. GC is calculated for each collection of binary images $(B_R, B_M, T_i(B_M)), i = 1, 2, 3, 4$, being in this example the mapping function T_4 chosen by our scheme.	55
Figure 16 – MRE for the registration methods when setting different geometric transformation models.	66
Figure 17 – Number of times that a certain geometric transformation model is chosen as optimal by our CG scheme and the MRE.	67
Figure 18 – Number of transformations selected by the GC from best (Rank 1) to worst (Rank 4). Clearly, the histograms show that GC tends to determine the best transformations for each method.	68
Figure 19 – Qualitative comparison for a given image pair when assessed by all the five evaluated methods.	69
Figure 20 – Color labeling-based inspection: the composition of two binary vessel images B_I (green) and B_R (magenta) so that complementary color channels are used to drive the visual analysis. A more substantial amount of white pixels means better registration outcomes.	69
Figure 21 – Visual results generated by VOTUS when applied to arbitrary images with different losses of quality from Dataset 1 (first two columns) while the remaining ones ticked with anatomical changes, which have been collected from FIRE.	71
Figure 22 – Qualitative comparisons from sample images collected from Dataset 1.	74
Figure 23 – Boxplots of the GC scores in Dataset 1 for the methods: VOTUS (■), H-M'17 (■), GDB-ICP (■), ED-DB-ICP (■), SIFT+WGTm (■), SURF+WGTm (■) and RIR-BS (■).	75
Figure 24 – MRE for a given image pair from Dataset 1.	75
Figure 25 – Boxplots of the MRE scores for VOTUS (■) and H-M'17 (■) in all datasets.	76
Figure 26 – Limitations of VOTUS: images pairs without the ONH as reference (top row) and with lack of focus (bottom row).	78
Figure 27 – VOTUS extended to multimodal fundus images: binary vessels overlapping image (left) and the classical checkerboard representation (right).	79

LIST OF TABLES

Table 1 – Image processing optimal transport research summary	33
Table 2 – Geometric processing optimal transport research summary	34
Table 3 – Graph matching comparison methods summary.	58
Table 4 – Image registration methods.	60
Table 5 – Number of images registered varying the features composing the cost matrix, having the best case in bold and the second best underlined	63
Table 6 – Comparative results using AUC for each graph matching method in our pipeline. Best results in bold.	64
Table 7 – Comparative results using AUC for different strategies of outlier removal in our pipeline. Best results in bold.	65
Table 8 – Statistics of the registration methods in their both standard and GC-optimized versions when assessed by the MRE measure.	65
Table 9 – Number of registered images for different parameters of the MRE for both standard and GC-driven registration methods.	66
Table 10 – Comparative results of VOTUS against other image registration methods using AUC. The best results and the second best are shown in bold/red and italic/blue, respectively.	72
Table 11 – The p -values of the registration results by the other methods in comparison with VOTUS.	73
Table 12 – Computational times (in seconds).	73

CONTENTS

1	INTRODUCTION	21
1.1	Optimal transport	24
1.2	Relevance and contributions	25
1.3	Dissertation outline	26
2	OPTIMAL TRANSPORT	27
2.1	Optimal transport problem	27
2.2	Optimal transport discretization	28
3	RELATED WORK	31
3.1	Optimal transport applied to image processing	31
3.2	Optimal transport applied to geometric processing	33
3.3	Fundus image registration: a review	34
4	REGISTRATION BY OPTIMAL TRANSPORT	39
4.1	Pipeline	39
4.2	Graph extraction	40
4.2.1	<i>Graph simplification</i>	41
4.3	Graph matching by optimal transportation	41
4.4	Ghost nodes for mass balance and outlier removal	43
4.5	Building the cost matrix	44
4.6	Feature extraction	45
4.7	Outlier filtering	47
4.7.1	<i>RANSAC</i>	47
4.7.2	<i>WGTM</i>	48
4.7.3	<i>DeSAC</i>	49
5	REVISITING FUNDUS IMAGE TRANSFORMATIONS	51
5.1	Transformation models	52
5.2	Best fit transformation	53
5.2.1	<i>Gain Coefficient</i>	54
6	RESULTS, COMPARISONS, AND EVALUATION	57
6.1	Graph matching methods	57

6.2	Registration methods	59
6.3	Metrics	61
6.4	Datasets	61
6.5	Pipeline steps evaluation: a quantitative assessment	62
6.5.1	<i>Feature selection</i>	63
6.5.2	<i>Graph matching comparison</i>	64
6.5.3	<i>Outlier removal comparison</i>	64
6.5.4	<i>Best fit transformation results</i>	65
6.5.5	<i>Image registration comparison</i>	68
6.5.6	<i>Qualitative analysis</i>	68
6.5.7	<i>Quantitative analysis</i>	70
6.5.8	<i>Statistical hypothesis testing</i>	72
6.6	Computational aspects	73
7	CONCLUSIONS	77
7.1	Discussion and limitations	77
7.2	Conclusions and future work	78
7.3	Bibliographic production related to this dissertation	80
7.4	Other bibliographic production	80
	BIBLIOGRAPHY	83

INTRODUCTION

Due to the heterogeneity and volume of information collected by several branches of science, there is a wide variation of problems in the area of image processing. Among those, several problems can be solved through mappings and distribution comparisons, from which can be cited: image stitching, morphing, registration, segmentation, video stabilization, among others. Furthermore, one interesting field of image processing that has become more popular over the years is the medical image processing one. With the advance of digital medical equipments with several examinations that are being done by visual analysis using images, from MRIs to CT-scans, more and more the medical field is closer to computer science. From the image processing field we will focus on the eye fundus image processing, which is also one of the recent popular fields of medical image processing (POPLIN *et al.*, 2018). An example of these images can be seen in Figure 1.

Figure 1 – Examples of eye fundus images.

In 2012, the World Health Organization (WHO) estimated that 39 million people in the world are blind, 285 million are visually impaired and 246 million have low vision degree (World Health Organization, 2014). Considering all these disorders, *glaucoma*, a serious disease that affects the eyes is considered the second leading cause of blindness worldwide (QUIGLEY; BROMAN, 2006). According to the American Academy of Ophthalmology, glaucoma is a complicated condition that damages eye's optic nerve. It regularly occurs when a fluid (called

aqueous) builds up in the front part of the eye, increasing the pressure on it. In general, the glaucoma pathology can be broadly classified into two core types, namely “open-angle” and “closed-angle” (or “angle closure”), both of them described in terms of the angle delimited between the iris and eye cornea. The open-angle case, more drastic, appears suddenly leading to the sudden vision loss while the closed-angle tends to progressively advance at a slower rate. As the medical diagnosis is mostly accomplished by human inquiry for glaucoma and other eye disorders, the use of image processing algorithms became a necessity (LALIBERTÉ; GAGNON; SHENG, 2003; KNOW *et al.*, 2009), especially when an ophthalmologist has to manage a large set of fundus images. Such computing progress has paved the way for clinicians and medical specialists to cover more patients as well as to reach greater diagnostic accuracy. However, in practice, the medical inspection of retinal images is still manually performed, in an attempt to carefully identify and track the progress of eye diseases. Moreover, these visual inspections are quite time-demanding, besides they also depend on the physician experience in order to succeed (KARALI *et al.*, 2004), noticing that some pathologies can require intensive examination over many years to be effectively detected.

The process of mapping a given image onto another image of the same scene is popularly known as *image registration* (VIERGEVER *et al.*, 2016). This procedure is commonly applied by many medical imaging systems where images captured at different times, scales and lighting conditions need to be carefully processed and aligned to ensure more consistency in clinical examinations. In fact, the efficacy in preventing eye’s diseases as well as to prescribe more comprehensive diagnostics highly depend on the success of those computer applications to perform registration accurately (SCHULTHEISS; TOMÉ; ORTON, 2012). Moreover, computer-aided medical technologies have also become a trend in the planning of optical surgeries (ONG *et al.*, 2015).

Regarding imaging solutions, there are a variety of computational approaches to support eye diseases detection (ZHANG *et al.*, 2014; RAVUDU; JAIN; KUNDA, 2012; KAUPPI, 2010). Among those, not many are directly aimed to register pairs of retinal images under real-world conditions as noise from deformations in the eye fundus that may occur due to diseases as the ones presented in Figure 2.

As physicians capture a set of images from the same patient over time, techniques devoted to registering retinal images help these experts to identify and monitor any signs of disease manifestation. Although existing techniques produce satisfactory results under certain conditions, they can deform the sequence of fundus images in many pragmatic situations. For instance, if the overlap between two images is not so apparent, the relative changes in the images become computationally unfeasible to be detected (DENG *et al.*, 2010). Additionally, retinal images often present textureless regions that are irregularly illuminated, hampering the correct matching between the blood vessels of the captured images.

Despite the benefits and advances promoted by computer-guided diagnostics, one must

Figure 2 – Examples of eye fundus images pairs, as columns, with noise from patients with disease-related deformations and illumination artifacts.

observe that the methods used to efficiently compute the registration models on fundus images can produce inaccurate outputs in many pragmatic cases, or even totally fail in this task, thus reducing the effective performance of several automatized registration systems, as previously reported by Philip, Cowie and Olson (2005), Abramoff *et al.* (2008), Pieczynski and Grzybowski (2015). To properly attempt overcoming this problem, a tremendous number of image registration techniques have been proposed in the last few decades, which vary in terms of many different resources (SOTIRAS; DAVATZIKOS; PARAGIOS, 2013; SCHNABEL *et al.*, 2016; VIERGEVER *et al.*, 2016; ZITOVÁ; FLUSSER, 2003; CATTIN *et al.*, 2006; YANG *et al.*, 2007c; RUECKERT; ALJABAR, 2010; TSAI *et al.*, 2010a; M; MOSES, 2014). Metrics for accuracy validation have also been considered in this context (LEE *et al.*, 2010; VARADHAN *et al.*, 2013; EL-GAMAL; ELMOGY; ATWAN, 2016), in an effort to check the real impact of existing registration tools while still making sure that clinical requirements will be strictly satisfied with those tools.

With this context in mind, in this research we will take advantage of the mathematical foundation of the theory of *Optimal Transport* (OT) (SANTAMBROGIO, 2015; SOLOMON, 2018), introduced bellow, as a mean to solve the fundus image registration problem.

1.1 Optimal transport

In 1781, the mathematician Gaspar Monge published his famous problem (VILLANI, 2009a) described as follows: “Assuming that a certain amount of sand must be conveyed from the soil to construction sites, where input and output locations are known, what is the best way to perform such a transportation?”. In summary, we need to discover the optimal map that assigns to the construction sites portions of sand coming from different resources. Considering that the material transportation has a cost, the optimal map should have a minimum total cost. Notice that the answer to this problem can be useful to many areas of knowledge and with this analogy in mind, next we describe this problem formally.

Several computational problems can be treated as optimal transport problems, where the objective is to find a mapping that minimizes the transportation cost, or in other words, to find the minimal total distance between two sets of samples. Figure 3 illustrates graphically the optimal transport problem, solved using linear programming, where the red set is mapped to the blue set according to optimal criteria. More specifically the objective was to find the set correspondences where the sum of distances between the matched points of one set to the other was minimal.

Figure 3 – 2D Optimal transport problem.

These theories and their extensions are used in several areas from physics to image processing (VILLANI, 2009b). Recently, is observed an increase in the number of researches applying optimal transport to geometric processing (RABIN; PEYRÉ; COHEN, 2010; DOMINITZ; TANNENBAUM, 2010; GOES *et al.*, 2011) and image processing (RUBNER; TOMASI; GUIBAS, 2000; HAKER *et al.*, 2004a; GOES *et al.*, 2012; FERRADANS *et al.*, 2013; WANG *et al.*, 2013), including some medical applications (REHMAN *et al.*, 2009a; ZHAO *et al.*, 2013).

1.2 Relevance and contributions

The proposed method, which we call **VOTUS** (*Vessel Optimal Transport for alignment of retinal fundUS images*), bears many attractive functionalities such as registration accuracy (even for images with a high level of specular noise), good performance for high-resolution images, and consistent solutions for the registration problem. Our method is stable, once achieved an acceptable result for a given image pair, this result remains unaltered even running VOTUS several times for the same image pair, due to its deterministic nature. Moreover, VOTUS registration may enhance the perception of anatomy changes between the images, with the objective of aiding health specialists to diagnose eye diseases and other complex impairments. A comprehensive set of comparisons and evaluations along the state-of-the-art image registration methods on real medical datasets, which attests the good performance of our approach. In summary, our main contributions are:

- A novel automatic framework which relies on OT theory to precisely match and align vessels extracted from retinal fundus image pairs;
- A customized OT cost function for blood vessels matching based on dissimilarity measures between fundus image features;
- An improved extension of the DeSAC filtering ([MCILROY *et al.*, 2010](#)) which reuses our OT cost matrix, becoming very robust against outliers (noise) and well suited for our application;
- A simple fundus image feature that is invariant under translation and scale transformations using saliency cues.
- During this research, we also revisited the choice of the transformation usually applied to fundus image registration and our study demonstrates that no transformation is optimal for fundus registration: the ideal mapping varies according to the retinal image pair.
- A new metric that gauges the overlapping between fundus images, called here as *Gain Coefficient*.
- An effective methodology to select the best transformation candidate for the registration task that is simple-to-code and user-independent, has low computational cost and improves the accuracy of several popular image registration methods.

The technique described in this research presents a novel automated optimal transport-based framework which addresses the problem of fundus image registration, possibly aiding medical specialists in tracking potential eye degenerations and anatomy changes. Moreover, our approach ensures many attractive characteristics such as a deterministic solution for the matching issue, is

efficient as it can deal with high-resolution fundus images within feasible processing time, and promotes high accuracy in terms of registration quality, even for images containing degradations, such as specular noise and other potential artifacts. In addition to that, some of the steps adopted in our pipeline can be easily adopted by other registration methods, and we prove by example that this can influence positively in the accuracy of other commonly used registration methods.

1.3 Dissertation outline

This research will be divided as follows: in [Chapter 2](#) we first present the theory of optimal transport as it will be the mathematical basis of our research, in [Chapter 3](#) are presented the related works of optimal transport applications and fundus image registration; the developed registration method is shown in [Chapter 4](#); next in [Chapter 5](#) we review and provide our contribution to the last step of the fundus registration algorithms, regarding the geometric transformation model to be applied; the results are shown in [Chapter 6](#) and in [Chapter 7](#) are presented our conclusions.

OPTIMAL TRANSPORT

In this Chapter, we present the theory of the optimal transport problem and approaches to solve it ([VILLANI, 2003](#)).

2.1 Optimal transport problem

In 1781, Monge published his famous problem described as: assuming that there is a certain amount of dirt to be extracted from the soil and transported to other places where it should be used in constructions. The extraction sites and the final transportation sites are known. But the following assignment must be determined: to which site each amount of soil must be transported? Considering that the soil transport has a cost and that we should find the minimum total cost, this problem becomes important for several fields.

Observing that the response may be useful to several areas and maintaining this analogy in mind, below we define this problem formally.

First, we define a probability measure M as a real valued function over a space E , that satisfy the Kolmogorov axioms:

1. $M(e) \geq 0$ and $M(e) \in \mathbb{R}, \forall e \in E$;
2. $M(E) = 1$;
3. Let $\{A_i \mid i \in \mathbb{I}\}$ be a disjoint countable set in E , then $M(\bigcup_{i \in \mathbb{I}} A_i) = \sum_{i \in \mathbb{I}} M(A_i)$.

Then we can model the soil pile and the sites using two probabilities measures μ and ν , respectively, defined on measure spaces X and Y . We will assume that both the soil pile and the amount needed have the same volume, and by simplicity, for now, say this mass is 1. Then, let A and B be measurable subsets of X and Y , then we have that $\mu(A)$ and $\nu(B)$ are the amounts of soil in A and B . Let $c(x, y)$ be the cost to transport a unit of soil from a location $x \in X$ to $y \in Y$,

where c must be measurable and nonnegative. More formally, c should be a measurable map from $X \times Y$ to $\mathbb{R}_+ \cup \{+\infty\}$, since c can also assume infinite value.

A simplistic illustration of the transport problem is shown in [Figure 4](#).

Figure 4 – Illustration of the optimal transport problem.

Kantorovich was the mathematician who studied Monge's problem and received a Nobel Prize in economics for his work. Next we will study his approach.

To solve the optimal transport problem we will define a *transportation plan* by the probability measures π on the product space $X \times Y$. The amount of mass transported from x to y is defined as $d\pi(x, y)$ and for any plan π to be valid it needs to preserve the problem mass. Hence all the mass taken from x should coincide with $d\mu(x)$ and all mass transported to y needs to coincide with $d\nu(y)$, hence denote the set of valid transportation plans as $\Pi(\mu, \nu)$. Formally

$$\int_X d\pi(x, y) = d\nu(y), \quad \int_Y d\pi(x, y) = d\mu(x),$$

or in other words

$$\pi[A \times Y] = \mu[A], \quad \pi[B \times X] = \nu[B]$$

This mathematical definition of this transport problem is known as the *Kantorovich's optimal transport problem*:

$$\text{minimize } f(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y), \text{ for } \pi \in \Pi(\mu, \nu).$$

Also, $f(\pi)$ is the total transportation cost and the optimal transport plan is that minimize $f(\pi)$

$$T_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} f(\pi).$$

Kantorovich is, in fact, a relaxed version of the original Monge's problem because in the original version, no mass should be split so each x should be the transport to a single location y .

2.2 Optimal transport discretization

Since this problem is modeled over probabilities measures it can assume several configurations: they could be discrete distributions or continuous and of any dimensionality. The

optimal transport could be used even among distinct domains, where you can transport from continuous to discrete and vice-versa.

One possibility for above-described formulation could be that ν is a Dirac mass $\nu = \delta_a$. The Dirac mass at some point x will be $\delta_x : \delta_x[D] = 1$ if $x \in D$ and 0 otherwise. Hence

$$T_c(\mu, \delta_a) = \int_X c(x, a) d\mu(x),$$

then there is a single element in $\Pi(\mu, \nu)$ and all the mass of μ should be transported to this element.

Another example could be taking X and Y as discrete spaces, where the points have all the same mass, for simplicity,

$$\mu = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \quad \text{and} \quad \nu = \frac{1}{n} \sum_{i=1}^n \delta_{y_i}$$

then any measure in $\Pi(\mu, \nu)$ can be represented by an $n \times n$ bistochastic matrix $\pi = (\pi_{ij})_{i,j}$ meaning that all π_{ij} are nonnegative, the cost function can be also represented as an $n \times n$ matrix C and

$$\begin{cases} \forall x_i \in X, \sum_j \pi_{ij} = 1 \\ \forall y_j \in Y, \sum_i \pi_{ij} = 1. \end{cases} \quad (2.1)$$

The discrete optimal transport problem is then rearranged as follows:

$$\text{minimize } f(\pi) = \frac{1}{n} \sum_{i,j} C(x_i, y_j) \pi_{ij}, \text{ for } \pi \in \Pi', \quad (2.2)$$

where Π' holds all transportation plans that satisfy [Equation 2.1](#). When the cost $C(x_i, y_j)$ is chosen as $\|x_i - y_j\|^p$, the problem (2.2) is known as the p -Wasserstein distance ([SOLOMON, 2018](#)). This same formulation can be easily extended for points with different mass, as long as the mass is preserved. In [Figure 5](#) we show a few examples of general distributions that could be used.

Therefore, the optimal transport problem in the discrete form [Equation 2.2](#) can be successfully applied to our graph matching purpose, once we have a finite number of blood vessel nodes obtained from the fundus images that will be connected each other. In the following, we detail this process.

Although the Monge's problem has a solution by Kantorovich's discretization, the solution of a problem with n^2 variables has a high computational cost. One way to get around this problem, adopted in several papers seen in this chapter was sampling. A solution using sampling was also used in this research, seen in [section 4.2](#).

Most recently, Optimal Transport (OT) theory has figured as an effective and powerful tool to address sophisticated problems in various fields of study, as specific optimization functions can be applied to handle the problem as well as the solid theoretical background it

Figure 5 – Illustration of some distinct distributions that can be used in an optimal transport problem: discrete distributions where $\nu = a_i \delta_{x_i}$, from which uniform are blue and $a_i = \frac{1}{n}$ and arbitrary are in green where the area/length can be interpreted as a_i ; in red can be seen continuous distributions where $d\nu(x) = \rho(x)dx$.

relies on. Indeed, OT theory has served as the basis for issues in different contexts of applications including *Economic Theory* (CHIAPPORI; MCCANN; NESHEIM, 2010), *Geometry Processing* (SOLOMON *et al.*, 2015a) and *Pattern Analysis* (PAPADAKIS, 2015).

Despite its effectiveness and high capability in modeling complex problems, only a few OT-inspired approaches have been properly devised to address medical imaging applications. In this context, Haker *et al.* (2004b), Rehman *et al.* (2009b) apply OT concepts to formulate a first order PDE that determines a deformation map (a grid) used to align images of specific categories (e.g. sky and MIR images). Another interesting work is the survey presented by Papadakis (2015), which provides a comprehensive overview of recent OT-based methods for general image processing purposes.

Next, we apply the concepts of optimal transport to the problem of eye fundus image registration.

RELATED WORK

Optimal transport have several applications in both image processing and geometric processing. Next we present a review of these related works in order to better to understand the application of this theory and how it was adapted to each context as the advances in these areas can be used as the base for the proposed research. We will also take the liberty to use the term *optimal transport* to refer to the theory of optimal transport.

3.1 Optimal transport applied to image processing

In their research [Rubner, Tomasi and Guibas \(2000\)](#) created the measure called *Earth Mover's Distance* (EMD). EMD has been based on the problem of optimal transport, where the cost to transport one unit of mass is given by an $N \times M$ matrix. This metric has become popular for its versatility and robustness. Rubner, Tomasi and Guibas used EMD to image retrieval taking advantage of the *ground distance* to measure different types of image features of the images.

Given continuity to the EMD research [Ling and Okada \(2007\)](#) proposed a new version of the original EMD with improved efficiency (the complexity goes from $O(N^2)$ to $O(N)$), called EMD- L_1 where the L_1 norm is used as *ground distance*. They also proved that EMD- L_1 is formally equivalent to EMD. This research deals with two problems: the correspondence of interesting points and the form recognition. Later, to improving the computational effort of this technique, [Pele and Werman \(2009\)](#) created a new variant of EMD, which is more efficient, using thresholded distances (source code available online ¹). This method is compared against the previously cited methods EMD- L_1 and EMD. It is not only faster (or equally fast), but it also delivers better qualitative results for most image retrieval cases.

Among the researches using the Monge-Kantorovich formulation, [Haker et al. \(2004a\)](#) can be cited for applying the methodology to image registration and warping, where the trans-

¹ <http://www.ariel.ac.il/sites/ofirpele/FastEMD/>

ported mass is the image intensity. Their approach can also be applied to 3D data. However, according to the authors, their optimal transport will not be suitable to circumstances where the image intensity is not preserved in both image pairs, as the mass preserving assumption will become invalid and clearly, by concept, this is the case in fundus image registration, where images pairs may have different occlusions, illumination (time-lapse between exams) and diseases related artifacts.

[Rehman et al. \(2009a\)](#) also using the Monge–Kantorovich theory, apply the optimal transport formulation to 3D medical images non-rigid registration, as the ones obtained using magnetic resonance imaging (MRI). Using graphical cards they obtained meaningful *speedups*, considering the computational cost of solving a transportation problem. Also, using a grid that represents the resulting mapping one can visualize the deformation among the MRIs and, hence, provide meaningful postoperative information to physicians. Also, regarding the application of this technique to fundus registration the authors also state that image (volumes) pairs should have same total mass falling into the same problem as reported by [Haker et al. \(2004a\)](#). We solve this problem in our formulation in [section 4.4](#).

Using the discrete approach with *power diagrams*, where the *blue noise* is generated with arbitrary density functions, [Goes et al. \(2012\)](#) modeled their problem using the optimal transport formulation to create sets of high-quality blue noise points. [Goes et al. \(2012\)](#) achieved spectral and spacial improvements using the minimization with constraints that are natural to the optimal transport theory.

Another interesting research is from [Ferradans et al. \(2013\)](#). A discrete relaxation of the optimal transport problem to transfer the color pallets among images, what according to the authors, would not be possible in the traditional formulation. The mapping is not constrained to be bijective, and also needed to deal with two sets with differences in total mass (the color pallets). Hence, even when the problem does not comply with all the optimal transport restrictions, it is still possible to change the formulation by replacing restrictions (with all the mathematical formalism involved).

In their publication, [Wang et al. \(2013\)](#) present a framework to efficiently measure distances among pairs of large-scale images. The problem is modeled as a linear optimal transport problem and solved with linear programming. Also, the authors demonstrate an application where it is possible to visualize the variation between images.

[Table 1](#) summarizes the cited methods with interesting details. The column Sampling/Initialization describes the approaches to reduce computational costs, either by initializing the problem with a solution close to the optimal or sampling the problem to reduce the number of variables or in the case of EMD, changing the *ground distance* in order to reduce computational cost. Finally, we mention the method for solving the optimization or differential equation.

Table 1 – Image processing optimal transport research summary

Author/Year	Objective	Sampling/Initialization	Method/Solver
Rubner, Tomasi and Guibas (2000)	Image retrieval	Signatures	Transport simplex
Haker <i>et al.</i> (2004a)	Image registration	Initial mapping	Gradient descent
Ling and Okada (2007)	Histogram comparison	L_1 ground distance	Tree-EMD
Pele and Werman (2009)	EMD extension	Thresholded distances	Minimum path
Rehman <i>et al.</i> (2009a)	3D registration	Regular grid	Conjugate gradient
Goes <i>et al.</i> (2012)	Blue noise generation	Power diagrams	Gradient descent
Ferradans <i>et al.</i> (2013)	Discrete OT	Clustering	Primal-dual
Wang <i>et al.</i> (2013)	Image distances	Clustering	Linear programming

3.2 Optimal transport applied to geometric processing

In the context of geometric processing Liu *et al.* (2010) presents a learning method to find a proper *ground distance* to the cited EMD in the context of partial shape retrieval. In this research, EMD is converted to the dual problem in order to emphasize the similarities between the sample's features. Solomon *et al.* (2014) proposed a new way to efficiently compute the distance among points in discrete meshes based on EMD. They use a differential EMD finite element formulation. In their results, this new measure presents good qualities when compared with metrics frequently used in meshes, such as the biharmonic distance, and geodesic.

Using the Monge-Kantorovich approach Rabin, Peyré and Cohen (2010) perform the recovery of 2D and 3D geometric shapes, using the Wasserstein metric collected from geodesic data. Quantiles are computed from local geodesic to make a global descriptor, invariant to isometric transformations. Next, the Wasserstein-2 approximated distances are computed using an interactive method so that this distance will be minimal between two sets in $\mathbb{R}^N$. It should be noted that, a metric was created to take advantage of the specific data of the problem domain, an aspect that should be considered when applying the optimal transport theory in any context. Also, using the proprieties of conservation of mass in Monge-Kantorovich Dominitz and Tannenbaum (2010) presents a new technique to texture mapping, that preserves the area during mapping. In fact, the optimal mapping preserves measures, hence volume or area are naturally preserved.

Goes *et al.* (2011) proposed a new method of 2D shape reconstruction given as input the set of points with noise and *outliers*. In the author's words, the use of optimal transport brings several benefits, including a unified treatment of noise, *outliers*, edges and sharp features.

Later, Zhao *et al.* (2013) presented a new mapping method to surfaces with texture to the plane. The mapping is done using optimal transport while preserving the area. Unlike previous methods Zhao *et al.* (2013) used the Monge-Brenier theory and solved the problem using the Newton method.

Digne *et al.* (2014) devised a simplified sampling-based point cloud mesh reconstruction method. Instead of using the full set of points, they cut computational costs by sampling points randomly, followed by triangulation. A simplified mesh will be the result of this optimal transport

algorithm: using optimal transport: attractively points are relocated to obtain lower cost solved using linear programming to find the local minima, a relaxation of the optimal transport problem. Finally, filters are made to remove simplexes using a mass threshold given by the user.

Published in one of the most relevant conferences in geometric processing [Solomon *et al.* \(2015b\)](#) presented a new method to solve optimal transport problems in the geometric domains. After the domain discretization using the finite element method, convolutions between vectors and diffusion kernels are made iteratively. The iterative process has linear convergence which has low computational cost. Among the presented applications are: shape interpolation, mesh skeleton computation, lighting, color pallet transference in images, image interpolation. This is an indication of the wide variety an optimal transportation research can provide.

[Table 2](#) present a summary of the geometric processing papers presented, column Sampling/Initialization display the sampling or initialization process used to lower the computational cost of solving the optimal transport, in the last column the method used to solve the proposed problem is shown.

Table 2 – Geometric processing optimal transport research summary

Author/Year	Objective	Sampling/Initialization	Method/Solver
Liu <i>et al.</i> (2010)	Partial 3D retrieval	uniform sampling	Adaboost + EMD
Rabin, Peyré and Cohen (2010)	Shape retrieval	<i>Farthest point sampling</i>	Gradient descent
Dominitz and Tannenbaum (2010)	Texture mapping	Conformal map. - sphere	Moser (1965)
Goes <i>et al.</i> (2011)	2D reconstruction	2D Delaunay Triangul.	Interactive + greedy
Zhao <i>et al.</i> (2013)	Mapping to the plane	Conformal map. - plane	Newton
Digne <i>et al.</i> (2014)	3D reconstruction	3D Delaunay Triangul.	Linear prog.
Solomon <i>et al.</i> (2014)	Metric over surfaces	Finite element	ADMM
Solomon <i>et al.</i> (2015b)	geometric OT	<i>Heat Kernel</i>	Sinkhorn (1964)

For more applications of the optimal transport theory, we leave as reference the book by [Villani \(2009b\)](#). It is worth to mention that the author Cédric Villani received the prestigious *Fields Medal* due to his research on the subject.

As shown in the previous revision, approaches that use the optimal transport theory can make use of the mapping, restrictions over mass and minimize a custom measure relevant to the context of the problem, while naturally preserving a context-dependent property. It is also worth noticing that it is possible to solve the optimal transport problem in large-scale. Finally, it is concluded in this review that this mathematical formulation is a strong tool for solving problems in different contexts.

3.3 Fundus image registration: a review

Considering the vast amount of research on medical image registration, numerous approaches have been proposed over the years, especially for aligning images of the same modality as well as montage multimodal arrangements ([ZHOU *et al.*, 2014](#); [LEGG *et al.*, 2015](#)). In order

to better contextualize our work, we focus our discussion on techniques devoted to aligning images from a unique imagery resource. The intent to be a comprehensive survey of all existing methods is beyond the scope of this dissertation.

We group retinal fundus image registration methods in three main categories: *vessel-based*, *intensity-based*, and *keypoint-based*. The first category, vessel-based, relies on graph-based representations of the eye's retina to accomplish the registration, while the intensity-based one performs the eye's alignment by exploiting specific attributes of the images such as colors, contrasts, and gradients. Finally, the last group aims at establishing a correspondence between keypoints extracted from the images in an effort to map their more intrinsic characteristics.

Vessel-based techniques attempt to generate a tree-like model representing the vascular structures of the retina so that the topological arrangements of these structures are detected and captured. Techniques from this category typically comprise two main stages: *vessel extraction* and *vascular registration*. [Can et al. \(2002\)](#) accomplish both extraction and registration processes by combining curvature attributes of the retina surface with a motion eye predictive algorithm. [Chen et al. \(2011a\)](#) provided a structure-matching mechanism that encodes bifurcation areas located at the vessels and embeds into characteristic vectors the length and branching angles of these regions. By measuring the similarity between the vectors, they construct a transformation model which reflects into its formula the matching between the similar structures of the images. Unlike [Chen et al. \(2011a\)](#), an elastic scheme is introduced by [Fang and Tang \(2006\)](#) to reconstruct "vascular trees", where an optimization function is computed until the optimality matching convergence between rearranged and reference vascular models is achieved. Later, [Palraj and Venilla \(2013\)](#) formulate a hybrid approach for which the registration task is performed by promoting specific mutual information-based measurements and a simulated annealing scheme. [Deng et al. \(2010\)](#) proposed a graph matching method, using the topological structures and the shape of the vessels as feature points. Thus, the retinal image registration problem is converted into an edge-to-edge correspondence problem. Similar to [Deng et al. \(2010\)](#), [Chen et al. \(2017\)](#) presented a novel vascular image registration method based on network structure index and circuit simulation. Besides, [Serradell et al. \(2015\)](#) introduced a method for matching two- and three-dimensional graph structures extracted from microscopy image stacks, angiography data and retinal fundus images, which the authors called *Active Testing Search for Robust Graph Matching* (ATS-RGM).

Still, regarding vessel-driven algorithms, [Ghassabi, Shanbehzadeh and Mohammadzadeh \(2016\)](#) proposed an interesting structure-based registration method which is invariant under rotation and small-scale changes. In the first step, their method extracts enclosed regions by the vessels using watershed segmentation. After the segmentation, a boundary-based region descriptor computes the similarity between these regions in order to determine correspondences. Although the method provides good results for high-resolution retinal images, the performance of the method is compromised when facing images with large-scale differences.

Intensity-based methods encode brightness properties of the image such as gradients, color contrast, and image saturation into a cost function to proceed with the vessel alignment. They make use of classical optimization resources such as *simulated annealing* (RITTER *et al.*, 1999), *genetic algorithms* (MATSOPOULOS *et al.*, 1999; YANG; HUANG, 2008) and *particle swarm optimization* (PALRAJ; VENNILA, 2016) to find a global minimizer to the corresponding registration problem. The framework proposed by Zaslavskiy, Bach and Vert (2009) is a good representative of an intensity-based method, where several image processing filters are combined and a graph matching problem is solved to promote the definitive alignment. Despite the mathematical robustness provided by the optimization theory, an important issue with approaches purely based on the image intensity is the computational cost: finding the optimal solution may be a very time-consuming task, especially for high-resolution medical images. Adal *et al.* (2014) proposed a hierarchical retinal image registration method which combines intensity and vascular structural information. Firstly, an image normalization process is computed in order to reduce the effects of the local luminosity and contrast variations on the retina during the image acquisition. Then, the image alignment is performed using a vascular-weighted mean squared difference similarity measure between the images.

Keypoint matching-based methods face the image alignment problem through point cloud perspective. In this kind of approach, local invariant features are extracted from the images so that a correspondence between two sets of feature points is defined. For instance, Yang *et al.* (2007b) proposed a new algorithm called *General Dual Bootstrap Iterative Closest Point* (GDB-ICP). In each bootstrap region, GDB-ICP uses the well-known *Scale Invariant Feature Transform* (SIFT) descriptor (LOWE, 2004) to find initial matches and the alignment step is performed by two different types of keypoints: corner points and face points. However, GDB-ICP algorithm is not robust under high contrast variation in high-resolution retinal fundus images, as reported by Ghassabi *et al.* (2013). Improvements of the GDB-ICP algorithm were proposed by Tsai *et al.* (2010b), where an *edge-driven dual bootstrap ICP* (ED-DB-ICP) algorithm is employed so that vascular features taken from the retinal images are incorporated into SIFT keypoints. Chen *et al.* (2010) created a multimodal image registration method using a *partial intensity invariant feature descriptor* (PIIFD). This method detects the keypoints using the *Harris-Stephen corner detector* Harris and Stephens (1988), followed by the computation of the PIIFD at each point. Then, the PIIFDs are matched according to a nearest-neighbor criterion. However, the PIIFD framework is not able to detect scale invariant and repeated keypoints. Wang *et al.* (2015) presented a point matching algorithm for retinal image registration using a Gaussian model to remove outliers. Their method relies on PIIFD descriptor combined with a robust local feature detector called *Speeded-Up Robust Features* (SURF) (BAY; TUYTELAARS; GOOL, 2006). Hernandez-Matas, Zabulis and Argyros (2016), Hernandez-Matas *et al.* (2017c) build an interesting keypoint-based retinal image registration method that takes advantage of the eye fundus geometry by exploiting a spherical model representation of the eye and the relative pose estimation of the cameras that captured the images to be registered. Recently, Hernandez-Matas,

Zabulis and Argyros (2017) improved the accuracy of their previous method by using SIFT descriptor and bifurcations in order to extract robust keypoints. The main advantage of keypoint matching-based methods is their flexibility in handling many image classes, including ocular fundus images, but they often rely on random mechanisms such as *Random Sample Consensus* (RANSAC) (FISCHLER; BOLLES, 1981a) to be effective, producing non-deterministic solutions to the associated registration problem.

In this research, we present a new automatic registration framework to aid clinicians in identifying and tracking the evolution of eye diseases by aligning blood vessels in ocular fundus images. Our framework relies on the Optimal Transport theory (SANTAMBROGIO, 2015), a robust and versatile mathematical resource, and combines the flexibility of graph-based representation to precisely establish a matching between keypoints of the eye's blood vessels. The performance and usefulness of the proposed approach are verified by a comprehensive set of comparisons against competing and well-established image registration methods by using real medical datasets and a variety of measures typically employed as a benchmark by the medical imaging community.

In contrast to most existing OT-based algorithms, which generate deformation grids or deal with specific kind of image classes, our methodology provides a new facet of optimal transport towards aligning vascular vessels on the retinal images. Graph representatives are built from the pair of acquired images so that their nodes are viewed as a multi-valued set of features and evaluated as keypoints to achieve the registration. Once the keypoints are established, we compute their direct correspondence between the images by solving an optimal transport problem. Such a formulation allows us to accomplish the matching using a custom cost function that deals with outliers as long as the match is obtained. Furthermore, different from most existing image registration methods which compute non-deterministic solutions to match keypoints, our technique provides a definitive response to assign the correspondences, thus proceeding with the registration task without the necessity of estimating the most representative matching as output.

REGISTRATION BY OPTIMAL TRANSPORT

In this chapter we define the pipeline of our registration process and the details of each step that make possible to create a registration method based on the optimal transport theory which is accurate and robust to outliers.

4.1 Pipeline

As illustrated in [Figure 6](#), the VOTUS approach comprises four main steps, namely: *graph extraction and simplification*, *optimal node matching and outlier filtering*, *geometric transformation* and *registration*. First, we construct representative graphs for the vessels in the images so that most key attributes of the images (e.g., local shapes and topological relations) are captured and then embedded into these graphs. The graphs are used as input by our optimal transportation scheme aiming at precisely matching their sets of nodes. In the third stage of our method, potential outliers are automatically detected and, thus, removed from the node mapping generated during the previous step, while in the last stage, geometric transformations are finally applied between the images to effectively proceed with the registration. Details of each stage of the technique are presented below.

Figure 6 – Pipeline overview of VOTUS.

Figure 7 – Graph construction (from left to right): blood vessel binarization, blood vessel skeletonization, full graph after skeletization and its simplification drawn on the original images.

4.2 Graph extraction

Given two retinal fundus images I and R , the first step of our framework consists in generating graph-based representations of I and R so that the spatial and topological information of retinal veins are preserved as much as possible. In order to construct such representative graphs, firstly we use the technique proposed by [Bankhead et al. \(2012a\)](#) to create binary images, B_I and B_R . These binarizations reflect in their structures vessel image attributes such as lines, shapes and topological groupings. From B_I and B_R , we can generate skeleton images S_I and S_R by applying the iterative thinning method described by [Haralick and Shapiro \(2002\)](#), which shrinks large pixelated regions while still removing possible outliers.

Once the skeletons are created, one can build blood vessel representative graphs $\mathcal{G}_I(\mathcal{V}_I, \mathcal{E}_I)$ and $\mathcal{G}_R(\mathcal{V}_R, \mathcal{E}_R)$ by associating each segmented pixel of S_I (resp. S_R) to a node of $\mathcal{V}_I$ (resp. $\mathcal{V}_R$). The edges are defined by connecting two nodes $\mathbf{v}_I^i$ and $\mathbf{v}_I^j \in \mathcal{V}_I$, if the following equation is satisfied:

$$\left\| p_I^i - p_I^j \right\|_{\infty} \leq 1, \quad (4.1)$$

where p_I^i and p_I^j are the pixels of B_I assigned to the nodes $\mathbf{v}_I^i$ and $\mathbf{v}_I^j$, and the Chebyshev norm is

$$\|\mathbf{x}\|_{\infty} = \max\{|x_1|, \dots, |x_d|\} \quad (4.2)$$

for a vector $\mathbf{x}$ of d dimensions. The first three columns in [Figure 7](#) depicts the result of applying the aforementioned step to a pair of retinal fundus images.

4.2.1 Graph simplification

In order to take off duplicate information carried out by look-alike nodes in $\mathcal{G}_I$ and $\mathcal{G}_R$, we select the branch points of the skeleton images as keypoints, pulling out from the graphs any other internal node. As stated in Lee (2008), Chen *et al.* (2011a), these keypoints correspond to significant structures of the blood vessels. So, to accomplish this task, the nodes are progressively removed by a simple algorithm that covers three basic cases:

- if node holds one neighbor: simply remove it;
- if node holds two neighbors: merge neighbors and then remove it;
- if node holds more than two neighbors: merge all neighbors to the nearest keypoint node (via the Chebyshev norm in Equation 4.2) and then remove it.

Besides the ambiguity reduction, the reduced number of nodes alleviates the computational burden during the optimization process (section 4.3). Figure 7 (last two columns) illustrates this procedure showing the node reduction and the resemblance between the graphs for a pair of fundus images. Algorithm 1 explain in detail the simplification process.

4.3 Graph matching by optimal transportation

Let $\mathcal{G}_I$ and $\mathcal{G}_R$ be the blood vessel graphs computed in the previous stage of our framework. We formulate the graph matching process as an optimal transport problem that leads to “one-to-one” correspondence between graph nodes while still eliminating noticeable outliers. Our formulation is based on a linear program that solves an optimal assignment problem, an adapted case of the minimum cost flow problem (MAGNANTI; AHUJA; ORLIN, 2014). More specifically, we want to determine the optimal transportation plan $\mathbf{X}$, represented by a matrix of size $N_I \times N_R$, with $N_I = |\mathcal{V}_I|$ and $N_R = |\mathcal{V}_R|$, that matches the set of nodes $\mathcal{V}_I$ to $\mathcal{V}_R$. In other words, each coefficient X_{ij} of $\mathbf{X}$ provides the transfer route for conveying a certain quantity of mass from a node $\mathbf{v}_I^i \in \mathcal{V}_I$ to another one $\mathbf{v}_R^j \in \mathcal{V}_R$. $\mathbf{X}$ is a discrete form for π from Equation 2.2. Then the transportation plan $\mathbf{X}$ can be obtained by solving the following optimization problem:

$$\begin{aligned} & \arg \min_{\mathbf{X}} \sum_{i=1}^{N_I} \sum_{j=1}^{N_R} C_{ij} X_{ij} \\ & \text{subject to} \quad \begin{cases} \sum_j X_{ij} = \mu(\mathbf{v}_I^i), & \forall \mathbf{v}_I^i \in \mathcal{V}_I \\ \sum_i X_{ij} = v(\mathbf{v}_R^j), & \forall \mathbf{v}_R^j \in \mathcal{V}_R \\ 0 \leq X_{ij} \leq \beta \end{cases} \end{aligned} \quad (4.3)$$

where $\mu(\mathbf{v}_I^i)$ and $v(\mathbf{v}_R^j)$ are the masses located at the nodes $\mathbf{v}_I^i$ and $\mathbf{v}_R^j$, respectively, and the constant β is an upper bound. For simplicity, we choose $\mu(\mathbf{v}_I^i) = v(\mathbf{v}_R^j) = \beta = 1, \forall i, j$. Moreover,

Algorithm 1 – Graph simplification.

```

1: procedure SIMPLIFY_GRAPH(thinned binary image matrix  $\mathcal{B}$ ,  $\epsilon$ )
2:    $nodes \leftarrow allpoints(\mathcal{B})$ 
3:    $branch\_nodes \leftarrow branchpoints(\mathcal{B})$ 
4:    $edges \leftarrow chebyshev(nodes, nodes) \leq \epsilon$ 
5:    $nodes2remove \leftarrow$  index of the nodes not in  $branch\_nodes$ 
6:    $nodes2keep \leftarrow$  index of the nodes in  $branch\_nodes$ 
7:   while  $nodes2remove$  is not empty do
8:      $node_i \leftarrow nodes2remove(1)$ 
9:      $neighbors \leftarrow$  indexes where  $edges(node_i, :) == 1$ 
10:    if  $count(neighbors) == 2$  then
11:       $edges(neighbors(1), neighbors(2)) \leftarrow 1$ 
12:       $edges(neighbors(2), neighbors(1)) \leftarrow 1$ 
13:    else if  $count(neighbors) > 2$  then
14:       $index \leftarrow \arg \min chebyshev(nodes(node_i, :), branch\_nodes)$ 
15:      for  $i \in 1$  to  $count(neighbors)$  do
16:         $edges((neighbors(i), nodes2keep(index))) \leftarrow 1$ 
17:         $edges(nodes2keep(index), neighbors(i)) \leftarrow 1$ 
18:      end for
19:    end if
20:    Remove  $node_i$  from  $nodes2remove$ 
21:     $edges(:, node_i) \leftarrow 0$ 
22:     $edges(node_i, :) \leftarrow 0$ 
23:  end while
24:  remove from nodes all nodes with no edges
25:  return edges, nodes
26: end procedure

```

each value C_{ij} represents a coefficient of a *cost matrix* $\mathbf{C}$, which accounts for the transportation cost to move a portion of mass from $\mathbf{v}_I^i$ to $\mathbf{v}_R^j$. $\mathbf{C}$ will be defined in details on [section 4.5](#).

The set of linear constraints in [Equation 4.3](#) is imposed to manage the transport of mass between $\mathcal{G}_I$ and $\mathcal{G}_R$. More precisely, the mass located in $\mathbf{v}_I^i$ is allowed to be reconducted to some node $\mathbf{v}_R^j$, and vice versa, but assuring the system will preserve the total mass. Indeed, this optimization problem requires that $\mathcal{V}_I$ and $\mathcal{V}_R$ must have the same total mass, and since every single node is always placed as one mass unit, such a condition would imply that $N_I = N_R$. In practice, this situation rarely occurs so we must deal with this problem next.

One can notice that this formulation is the same as the discrete optimal transport with uniform mass presented in [Equation 2.2](#) with the restriction of [Equation 2.1](#). But our final formulation will have arbitrary discrete mass since we next we deal with mass difference among sets of nodes ($N_I \neq N_R$) and also outliers during the optimization step by adding our own restrictions to the original Kantorovich formulation.

4.4 Ghost nodes for mass balance and outlier removal

In order to equalize the mass difference between the graphs and also remove possible outliers, two *ghost nodes* $\mathbf{g}_I \in \mathcal{V}_I$ and $\mathbf{g}_R \in \mathcal{V}_R$ are incorporated into the optimization process. Without loss of generality, assuming that $N_I \geq N_R$, and the portions of mass in the ghost nodes $\mathbf{g}_I$ and $\mathbf{g}_R$ are given by α and $N_I - N_R + \alpha$, respectively. Quantity $N_I - N_R$ regulates the difference of masses between $\mathcal{G}_I$ and $\mathcal{G}_R$, while α is a fixed threshold which estimates the number of outliers in the mass transport model. Mathematically, the inclusion of the ghost nodes in the optimization problem means adding the following constraints into Equation 4.3:

$$\sum_{j=1}^{N_R} X_{N_I+1,j} \leq \alpha \quad \text{and} \quad \sum_{i=1}^{N_I} X_{i,N_R+1} \leq N_I - N_R + \alpha, \quad (4.4)$$

with $\alpha = \omega \min(N_I, N_R)$, $\omega \in [0, 1]$. The value ω depends on the number of outliers in the graphs, which varies from image to image. Regarding the computation of ω , we test a range of values and select the best estimative according to the gain registration measure presented in (MOTTA; CASACA; PAIVA, 2018).

Figure 8 – Optimal graph matching: the matrix $\mathbf{X}$ encodes the assignment between the nodes $\mathcal{N}_I$ (rows/orange) and $\mathcal{N}_R$ (columns/green), including their associated ghost nodes $\mathbf{g}_I$ (light orange) and $\mathbf{g}_R$ (light green).

Summarizing, the coefficients of the optimal transportation plan $\mathbf{X}$ provides the correspondences between the nodes $\mathcal{N}_I$ and $\mathcal{N}_R$ (see Figure 8). The additional constraints Equation 4.4 allow us to deal with outliers located at $\mathcal{G}_I$ and $\mathcal{G}_R$, as they will be attracted to the ghost nodes $\mathbf{g}_I$ and $\mathbf{g}_R$ carrying up to α unities of mass. The transportation cost between the ghost nodes is also required to be tuned into the model, and for this reason, we chose this cost as the minimum value of $\mathbf{C}$. Once the optimization is performed, all nodes matched with ghost nodes are interpreted as outliers and then discarded, i.e., $\mathbf{X}$ is updated by removing the rows and columns related to the ghost nodes. Thus, the set of correspondences between the matched nodes $\mathbf{v}_I^i$ and $\mathbf{v}_R^j$ is given by $\mathcal{M} = \{(i, j) \mid X_{ij} = 1\}$.

Figure 9 – Neighbor distance between the nodes $\mathbf{v}_I$ and $\mathbf{v}_R$.

In terms of computational implementation for the optimization problem in Equation 4.3, it can be efficiently solved by a mixed-integer linear programming solver (VIELMA, 2015).

4.5 Building the cost matrix

In this section, we will define our cost matrix which contains, for each combination node $\mathbf{v}_i \in \mathcal{V}_I$ and $\mathbf{v}_j \in \mathcal{V}_R$, the cost to transport one unit of mass from $\mathbf{v}_i$ to $\mathbf{v}_j$ as $C_{i,j}$. The cost matrix $\mathbf{C}$ encoded in Equation 4.3 is generated by summing matrices of node-to-node dissimilarities $\mathbf{D}^k$ and neighborhood-to-neighborhood dissimilarities $\mathbf{S}^k$, which are built according to a specified feature k of the image.

Each coefficient D_{ij}^k of the matrix $\mathbf{D}^k$ accounts for the dissimilarity between the nodes of $\mathcal{V}_I$ and $\mathcal{V}_R$. More specifically, given an image feature k , the cost to transport a unit of mass from the current node $\mathbf{v}_I^i$ to a second one $\mathbf{v}_R^j$, disregarding their neighborhoods, is computed by:

$$D_{ij}^k = \text{dis}_k \left(\text{desc}_k(\mathbf{v}_I^i), \text{desc}_k(\mathbf{v}_R^j) \right),$$

where desc_k denotes the feature descriptor k for the pixel associated with its corresponding graph node and dis_k represents the dissimilarity measure in the feature space k (e.g., negative cross-correlation, cosine distance or Euclidean distance).

In order to take advantage of the graph topology, we also embed the neighborhood of each node into the transport process using the adjacency information provided by the set of edges $\mathcal{E}_I$ and $\mathcal{E}_R$. The coefficients of $\mathbf{S}^k$ store the dissimilarity information between one-ring neighborhoods $\mathcal{N}_I^i$ and $\mathcal{N}_R^j$ w.r.t. $\mathbf{v}_I^i$ and $\mathbf{v}_R^j$, respectively. Although computing the dissimilarity between two sets is a task computationally expensive, we propose the Algorithm 2 as an efficient alternative to computing $\mathbf{S}^k$. Also, a graphic representation of the node to node neighborhood distance can be seen in Figure 9 where the neighbors of $\mathbf{v}_I$ are matched with $\mathbf{v}_R$'s while the node in yellow is marked to be penalized since it has no matching node. Our algorithm is “greedy” and eliminates the use of a more costly and complex procedure to measure the dissimilarity between

Algorithm 2 – Dissimilarity between neighborhoods.

```

1: procedure NEIGHBOR_DIST( $\mathcal{V}_I, \mathcal{V}_R$  and an image feature  $k$ )
2:   for  $i \in \{1, \dots, N_I\}$  do
3:     for  $j \in \{1, \dots, N_R\}$  do
4:       for  $\mathbf{v}_I^m \in \mathcal{N}_I^i$  do
5:         for  $\mathbf{v}_R^n \in \mathcal{N}_R^j$  do
6:            $f_I \leftarrow \text{desc}_k(\mathbf{v}_I^m)$ 
7:            $f_R \leftarrow \text{desc}_k(\mathbf{v}_R^n)$ 
8:            $\mathbf{F}(m, n) \leftarrow \text{dis}_k(f_I, f_R)$ 
9:         end for
10:      end for
11:       $D \leftarrow 0$ 
12:       $r \leftarrow |\mathcal{N}_I^i|, c \leftarrow |\mathcal{N}_R^j|$ 
13:       $N_{\min} \leftarrow \min(r, c), N_{\max} \leftarrow \max(r, c)$ 
14:      for  $p \in \{1, \dots, N_{\min}\}$  do
15:         $[\text{value}, \text{row}, \text{column}] \leftarrow \min(\mathbf{F})$ 
16:         $D \leftarrow D + \text{value}$ 
17:         $\mathbf{F}(\text{row}, 1 : c) \leftarrow \infty$ 
18:         $\mathbf{F}(1 : r, \text{column}) \leftarrow \infty$ 
19:      end for
20:       $\mathbf{S}^k(i, j) \leftarrow \frac{D}{N_{\min}} \left( 1 + \frac{N_{\max} - N_{\min}}{N_{\max}} \right)$ 
21:    end for
22:  end for
23:  return  $\mathbf{S}^k$ 
24: end procedure

```

sets. The resulting dissimilarity (line 20) is normalized by dividing the cumulative dissimilarity D by $N_{\min}$ while neighborhoods $\mathcal{N}_I^i$ and $\mathcal{N}_R^j$ with different cardinalities are penalized by the factor $(N_{\max} - N_{\min})/N_{\max}$, where $N_{\min} = \min\{|\mathcal{N}_I^i|, |\mathcal{N}_R^j|\}$ and $N_{\max} = \max\{|\mathcal{N}_I^i|, |\mathcal{N}_R^j|\}$.

Let M be the number of image features, the cost matrix is then defined by:

$$\mathbf{C} = \sum_{k=1}^M \mathbf{D}^k + \mathbf{S}^k, \quad (4.5)$$

where the elements of $\mathbf{D}^k$ and $\mathbf{S}^k$ are normalized in the range $[0, 1]$. Next, we describe in details the features and dissimilarity measures used to build the cost matrix shown on [Equation 4.5](#).

4.6 Feature extraction

In this step we discuss different types of image features used to achieve a context-relevant cost function. All features are computed in terms of the input ocular images and collected at the nodes or still from image patches around these nodes.

Node position is used in our approach as a positioning feature which aims at conveying geometrical information from the input graphs.

Figure 10 – Example of block features extracted in two pair images ordered in columns: first are shown the keypoints origin; the gray-image patch extracted is in the second column, and gabor filter patches in the last column.

Gray-image patches are directly collected from the gray-scaled version of the images and they hold information on the intensity of the keypoint neighborhood pixels and it is invariant to translation.

Gabor filter phases patches are used to extract texture information and time-frequency location from the given images. It provides robustness to our system even when varying the brightness/contrast content of the images as well. In our experiments, we set four Gabor filter-based features ranging from 0° to 90° with wavelength of 4 and 8 pixels/cycle. This feature is helpful when considering that we have no control over the illumination conditions changes between image pairs. A detailed survey of Gabor filters can be found in [Kamarainen, Kyrki and Kälviäinen \(2006\)](#).

Illustrated in [Figure 10](#) is an example of Gabor and gray patches extracted from a target keypoint. The patches extracted from the image pair visibly display a similar pattern since they were extracted from the same relative keypoint in both images.

Binary robust invariant scalable keypoints (BRISK) descriptor was used as a rotation invariant feature which gets binary strings of local intensity gradients with the goal of identifying the most relevant direction of each keypoint (nodes). For a more detailed discussion regarding BRISK descriptor, see [Leutenegger, Chli and Siegwart \(2011\)](#).

Scale Invariant Feature Transform (SIFT) is an image descriptor developed by [Lowe \(2004\)](#),

Figure 11 – Saliency-based vector (from left to right): input image, the most salient region (gray), its center point (in black), and vector field (yellow) built in terms of the ONH center point and the graph nodes.

which is invariant to translation, rotation, scale and robust to illumination changes. This descriptor was used for a large number of applications in computer vision and image processing.

Saliency-based vector is a novel, scale and translation invariant feature suited for fundus images. In order to define the most salient region, we adapt the saliency detection algorithm proposed by Zhu et al. (ZHU *et al.*, 2014) for the context of fundus image registration. After detecting the most salient region on the green channel of the image, for each node $\mathbf{v}$ a representative vector, which encodes the saliency cue at the node $\mathbf{v}$, is given by $\mathbf{u}_{\mathbf{v}} = \mathbf{v} - \mathbf{o}$, where $\mathbf{o}$ denotes the centroid of this region. In other words, a vector field is defined by the centroid of the most salient region and the graph nodes of the corresponding image, as shown in Figure 11. In fact, the most salient region tends to be located at the *Optic Nerve Head* (ONH) (HOOVER; GOLDBAUM, 2003).

The dissimilarity measure used in patch-based features (gray-images and Gabor filter) was the negative cross-correlation, while the Euclidean distance was applied in node-based features (BRISK, SIFT and node position), and the cosine distance was used in the saliency-based feature. From the set of features presented previously we tested all possible combinations to determine which is the optimal set for our problem. Results are presented in subsection 6.5.1.

4.7 Outlier filtering

Although the proposed method deals with outlier during the optimization, any remaining mismatch needs to be removed in order to maintain accuracy. In this section, three removal outlier approaches are investigated. We compare them and select the one which best suits our needs.

4.7.1 RANSAC

In order to deal with *outliers* that may occur after the optimization, the first method that was reviewed was RANSAC. Introduced by Fischler and Bolles (1981b) the RANSAC algorithm (RANdom Sample And Consensus) is an interactive method to estimate the parameters of a

given mathematical model. Among the advantages of this method, we can cite the fact that this method is robust to outliers and taking advantage of this quality we can use it to classify the correspondences obtained in anterior steps as inliers or outliers. RANSAC selects randomly a minimal sample of the data and fits the model, this set of data will grow iteratively in a way that the inliers are maintained while discarding the outliers. The algorithm stops when a number of iterations or, number of inliers is achieved.

The original RANSAC algorithm is composed of two steps that will be executed iteratively where a hypothesis is proposed and tested. Given a set of data $\mathbf{P}$, do for k iterations:

- *Hypothesis*: From a minimal set $\mathbf{P}_r$ of the data $\mathbf{P}$, randomly initialized, the model parameters are computed.
- *Test*: Next is tested which elements of the set $\mathbf{P}$ are consistent (inliers) with the model estimated. This set of inliers associated with the model is the consensus set.

Eventually, the model chosen is the one with more inliers. In our case, the mathematical model adopted is the 2D homography solved with the Direct Linear Transformation (DLT) algorithm presented by [Hartley and Zisserman \(2003\)](#).

The homography, also called the projective/perspective transformation is explained in [section 5.1](#).

In this case, an element of the set classified as an inlier or outlier is, in fact, two points in space, the correspondent points. Hence, given two sets of points $\mathbf{x}$ and $\mathbf{y}$ where x_i is the corresponding point of y_i , the i -th element will be classified as inlier according to the homography $\mathbf{H}$ (defined in [section 5.1](#)) in both directions using the symmetric transfer error ([HARTLEY; ZISSERMAN, 2003](#)) given by

$$STE_i = \sum (x_i - \mathbf{H}^{-1}y_i)^2 + \sum (y_i - \mathbf{H}x_i)^2.$$

A sample i is classified as inlier if STE_i is lower than a threshold t given by the user. Although this technique presents several advantages, an obvious disadvantage is its non-deterministic nature, due to the randomness of the initialization. To deal with that, we tested two possibilities instead: one is an approach called Weighted Graph Transformation Matching (WGTM) ([IZADI; SAEEDI, 2012](#)) and the other is based on the algorithm called DeSAC which we adapted to our context, these algorithms are presented next.

4.7.2 WGTM

WGTM (*Weighted Graph Transformation Matching*) ([IZADI; SAEEDI, 2012](#)) is a graph-based method, starts from an initial solution, from that constructs the graphs (by the K-Nearest Neighbors algorithm ([DUDA; HART; STORK, 2012](#))) and uses angular distances between edges

to iteratively remove outliers from matches. One good point to notice is that the angles are invariant to scale and rotation, a useful property in the context of graph matching. After removing such outliers the method re-evaluates the angles and stops according to two conditions, one is the maximum weight threshold ε given by the user and the other is if the mean weight is decreasing. For our comparisons, we used $K = 20$ and $\varepsilon = 0.001$ the best parameters that we found in the range suggested by the authors.

4.7.3 DeSAC

Aiming at removing possible outliers originated during the previous steps, we adapted the method proposed by [McIlroy et al. \(2010\)](#), known as DeSAC (*Deterministic Sample And Consensus*), for our specific application. Since DeSAC requires a matching probability value P_{ij} for each correspondence $(i, j) \in \mathcal{M}$, P_{ij} is given by a convex combination and parameterized in terms of the coefficients of the OT cost matrix $\mathbf{C}$, as follows:

$$P_{ij} = 0 \cdot t_{ij} + 1 \cdot (1 - t_{ij}) = 1 - t_{ij} \quad \text{with} \quad t_{ij} = \frac{C_{ij} - m}{M - m},$$

where M and m are the maximum and minimum value of the set $\{C_{ij} \mid (i, j) \in \mathcal{M}\}$, respectively. Note that, low transportation costs C_{ij} implies in high matching probabilities P_{ij} , and vice-versa. [Figure 12](#) shows the result of our improved DeSAC filtering on a pair of fundus images. The removal of these remaining outliers are indispensable to high-accuracy result, since the transformation will be estimated from the matched points. More details about the effectiveness of this approach and comparisons are presented in [subsection 6.5.3](#).

Finally, the registration is defined after outliers filtering by the geometric transformation scheme presented by [Motta, Casaca and Paiva \(2018\)](#) and described next.

Figure 12 – Outlier filtering: before filtering (yellow) with outliers (cyan) and after the improved DeSAC filtering (white).

REVISITING FUNDUS IMAGE TRANSFORMATIONS

While the plurality of the image registration methods has strongly contributed to the progress of general computer-aided applications over the years, an important issue not always covered by most existing registration methods is the choice of the best candidate for the geometric alignment function. Indeed, the methods assume as quadratic the transformation used to accomplish the alignment between the retinal images ([VIERGEVER *et al.*, 2016](#); [ZITOVÁ; FLUSSER, 2003](#)). But by taking quadratic/spherical transformations to register fundus images in the majority of cases, one could be discarding potential mapping candidates for which the registration accuracies are higher. Moreover, from a more suitable geometric transformation, one could be able to deal with several cases of unregistered images where classic alignment functions cannot succeed as expected.

Next, we investigate how to choose optimal transformations through a robust and user-independent decision metric which improves the accuracy performance of several state-of-the-art registration algorithms on eye fundus images. From a novel overlapping metric specifically designed to select the best image transformation, i.e., the one that maximizes the registration accuracy, we improve the performance of well-established registration methods while still maintaining their current computational costs. In more specific terms, the proposed metric consists of checking topological and geometric changes in the eye's blood vessels and it does not require for ground truth in practical usages were also computed in real-time even for large images commonly found in medical contexts, also serving as a useful index to support registration applications at interactive rates.

5.1 Transformation models

We consider four well-known transformation models in our analysis: *similarity*, *affine*, *homography* and *quadratic*. Notice that, although we take the above-mentioned geometric models as a baseline, the approach presented here can be applied for other sets of transformations as well. Next, we briefly describe each one of these transformations.

Points in the projective plane. Points in Euclidean plane can be represented in the *projective plane* $\mathbb{P}^2 = \mathbb{R}^3 \setminus (0,0,0)^\top$ using homogeneous coordinates $\tilde{\mathbf{x}} = (\tilde{x}, \tilde{y}, \tilde{w})^\top$. Moreover, a given point $\tilde{\mathbf{x}} \in \mathbb{P}^2$ can be transformed back to Euclidean plane by dividing through by the last entry, i.e.,

$$\tilde{\mathbf{x}} = (\tilde{x}, \tilde{y}, \tilde{w})^\top = \tilde{w} \left(\frac{\tilde{x}}{\tilde{w}}, \frac{\tilde{y}}{\tilde{w}}, 1 \right)^\top = \tilde{w}(x, y, 1)^\top = \tilde{w}\bar{\mathbf{x}},$$

where $\bar{\mathbf{x}} = (x, y, 1)^\top$ is the augmented vector of $\mathbf{x} = (x, y)^\top \in \mathbb{R}^2$.

Similarity transformation. This transformation can be written as $\mathbf{x}' = s\mathbf{R}_\theta \mathbf{x} + \mathbf{t}$, where s is a scale factor, $\mathbf{R}_\theta$ accounts for the rotation matrix by an angle θ , and $\mathbf{t} = (t_x, t_y)^\top$ determines the translation vector. The geometric mapping is properly computed as:

$$\mathbf{x}' = \begin{bmatrix} s \cos \theta & -s \sin \theta & t_x \\ s \sin \theta & s \cos \theta & t_y \end{bmatrix} \bar{\mathbf{x}}.$$

Affine transformation. In two-dimensional coordinate systems, the affine transformation is determined by $\mathbf{x}' = \mathbf{A}\bar{\mathbf{x}}$, where $\mathbf{A}$ is an arbitrary 2×3 augmented matrix. This matrix is known as the *affine transformation matrix*.

Projective transformation. Also called *perspective* or *homography mapping* (SZELISKI, 2010), this transformation is computed through homogeneous coordinates $\tilde{\mathbf{x}}' = \tilde{\mathbf{H}}\tilde{\mathbf{x}}$, where $\tilde{\mathbf{H}}$ is a 3×3 matrix. The resulting vector $\tilde{\mathbf{x}}'$ in homogeneous coordinates can be converted to $\mathbf{x}' = (x', y')^\top \in \mathbb{R}^2$ by:

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \quad \text{and} \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}.$$

Quadratic transformation. The quadratic transformation is a second order polynomial function which is defined as: $\mathbf{x}' = \mathbf{B}\mathbf{p}(\mathbf{x}) + \mathbf{A}\bar{\mathbf{x}}$, where $\mathbf{A}$ and $\mathbf{B}$ are 2×3 matrices, and $\mathbf{p}(\mathbf{x}) = (x^2, y^2, xy)^\top$.

Notice that the matrices that appear in the geometric transformations are defined so as to allow the use of a specific registration method. Considering some particular differences between these transformations, similarity-based preserves angles between lines, affine-based ensures that parallel lines will be kept after the geometric alignment, projective-based conserves straight lines and, finally, quadratic-based can embed lines into the curved shape of the retina, but it also

Figure 13 – Illustrative example of the different characteristics of the transformations considered.

introduces non-linear errors, as shown by [Eldeeb and Fahmy \(2014\)](#). A visual representation of these differences can be seen in [Figure 13](#).

The reader can observe that the only non-linear mapping considered in our analysis is the quadratic one. In fact, this is the most common choice when coping with fundus images ([VIERGEVER et al., 2016](#); [ZITOVÁ; FLUSSER, 2003](#)). Its use is endorsed by the spherical shape of the eye, but one may note that non-linear transformations can also lead to non-linear errors. For instance, the monitoring of some pathologies, including glaucoma, is accomplished from a sequence of eye's exams, so these errors can influence the diagnostic if they were not taken into account.

In the following, we present a simple methodology to determine the best geometric transformation for a given pair of fundus images.

5.2 Best fit transformation

In order to define the best transformation to each image pair we first revisited some classic image similarity metrics:

Mutual Information (MI) quantifies how much the images R and I share the same information taking into account their relative entropy, that is, it measures the degree of uncertainty contained in R (or I). MI is given by:

$$MI = \frac{H(R) + H(I) - H(R, I)}{\sqrt{H(R)H(I)}}, \quad (5.1)$$

where quantities $H(R, I)$, $H(R)$ and $H(I)$ denote the joint and marginal entropies of the images respectively ([GRAY, 2011](#); [STREHL; GHOSH, 2002](#)). In this metric, the closer the value is to 1, the better the registration.

Universal Image Quality Index (UIQI) is based on statistical measures in order to penalize three impairments on a pair of images: *correlation loss*, *luminance distortion* and *contrast blurring* ([WANG; BOVIK, 2002](#)). The contribution of each pair of registered image patches $x_j \in I$ and $y_j \in R$ is computed by

$$Q_j = \frac{4\sigma_{x_j y_j} \bar{x}_j \bar{y}_j}{(\sigma_{x_j}^2 + \sigma_{y_j}^2)(\bar{x}_j^2 + \bar{y}_j^2)}, \quad (5.2)$$

where $\sigma_{x_j y_j}$ represents the correlation coefficient between x_j and y_j , σ_{x_j} and σ_{y_j} are the standard deviation, and $\bar{x}_j$ and $\bar{y}_j$ are the means of x_j and y_j respectively. Finally, UIQI is obtained from the average of the values Q_j over the entire image. In this metric, the closer the value is to 1, the better the registration.

Structural Similarity Index (SSIM) is also based on contrast and luminance variation as the previous case but it introduces the concept of structural similarity to create a measure closer to human perception. Considering a pair of registered image patches $x_j \in I$ and $y_j \in R$, SSIM is given by an average of the following values:

$$S_j = \frac{(2\bar{x}_j\bar{y}_j + c_1^2 L^2)(2\sigma_{x_j y_j} + c_2^2 L^2)}{(\bar{x}_j^2 + \bar{y}_j^2 + c_1^2 L^2)(\sigma_{x_j}^2 + \sigma_{y_j}^2 + c_2^2 L^2)}, \quad (5.3)$$

with constants $c_1 = 10^{-2}$ and $c_2 = 3 \times 10^{-2}$. The value L is the dynamic range of the pixel values as used by Wang *et al.* (2004). In this metric, the greater value means better registration.

5.2.1 Gain Coefficient

After this revision and some practical tests we realized that these measures are all affected by the number of black pixels, characteristic in fundus images, that will not be registered and then would be penalized on all these metrics. So we proceed by creating our own metric based on the segmented vessels and a simplistic comparison (a logical *and*) to check for registration quality.

So, let $\mathcal{T} = \{T_1, \dots, T_n\}$ be the set of possible transformations, to choose the best geometric mapping which maximizes the accuracy for a given registration method, we apply the following sequence of steps:

- S.1** Extract the blood vessels binary images B_R and B_M from the fundus image pair I_R and I_M respectively (for more details, see Bankhead *et al.* (2012b));
- S.2** Compute each transformation i (if possible) to the image to be mapped, B_M , generating a set of transformed images $T_i(B_M)$;
- S.3** For each collection of images $(B_R, B_M, T_i(B_M))$, compute the registration accuracy based on a new overlapping metric called *Gain Coefficient* (GC);
- S.4** The optimal geometric transformation is the one that maximizes the GC.

The rationale behind our approach is quite simple: no exclusive transformation is optimal, i.e. the mapping depends on the image pair as well as the matching points given by the registration method. For this purpose, we maximize the GC for all the available transformations, which is computationally inexpensive and can be computed “on-the-fly”. Finally, the proposed GC metric that quantifies the registration accuracy for a given transformation T is computed as follows:

Figure 14 – Measuring the registration accuracy from the GC metric. First, blood vessels B_R and B_M are built from the fundus images I_R and I_M . Then, a geometric transformation model T is applied on B_M so as to generate $T(B_M)$. Finally, the Gain Coefficient is computed as in Equation 5.4.

Figure 15 – Illustration of our GC-based approach for selecting the optimal transformation. GC is calculated for each collection of binary images $(B_R, B_M, T_i(B_M)), i = 1, 2, 3, 4$, being in this example the mapping function T_4 chosen by our scheme.

$$GC(B_R, B_M, T) = \frac{|B_R \cap T(B_M)|}{|B_R \cap B_M|}. \quad (5.4)$$

Equation 5.4 states that if a geometric transformation aligns more pixels after the mapping, GC will be greater than 1, as the numerator indicates the amount of overlapped pixels on B_R and $T(B_M)$, while $B_R \cap B_M$ quantifies the input overlap wherein no transformation is applied. Figure 14 illustrates how the GC properly works while Figure 15 shows an illustrative case covering the steps S.4 and S.5 of our selection criterion. Notice from the graphical sketches that we depict composed images where B_R is set in the green channel, B_M (or $T(B_M)$) is represented in both red and blue channels (magenta), and finally when these images completely overlap we have the white color.

The intuition is that if a transformation performs towards more overlapping between the image pair this coefficient will be greater than zero and a value lower than 1 indicates that the registration has less overlapping than the starting input hence being unsuccessful.

RESULTS, COMPARISONS, AND EVALUATION

In this section, we present an extensive set of comparisons to better assess our technique, beginning with a comparative study of our pipeline against state-of-the-art methods along with classical methods frequently used in practical applications. To that, we consider three representative datasets that address various pragmatic situations ranging from the presence of specular noise to the diagnosis of eye diseases.

First, we must introduce in [section 6.1](#) the graph matching methods compared to our own. In [section 6.2](#) we present the image registration methods, the metrics used to summarize our results. The process to compute them are in [section 6.3](#). In [section 6.5](#) we show these comparisons first by assessing our feature choices in [subsection 6.5.1](#) and then our outlier removal method in [subsection 6.5.3](#). In [subsection 6.5.5](#) we effectively evaluate our registration against other methods and in [subsection 6.5.8](#) we present the statistics results for the previous registrations per dataset. Finally, we cover the implementation details in [section 6.6](#) along with a comparison of the computation times.

6.1 Graph matching methods

Next will be described several graph matching methods that will be compared with our optimal transport based approach. We will compare these methods in an isolated manner with our matching method in order to evaluate the quality of our graph matching in the context of fundus image registration.

Graduated assignment (GA) [Gold and Rangarajan \(1996a\)](#) perform an assignment optimization by starting with an initial value and then solving a sequence of convex approximations, by iteratively maximizing a Taylor expansion and controlling the accuracy by a control parameter

β (KOSOWSKY; YUILLE, 1994) that is updated every iteration.

Probabilistic matching (PM) Zass and Shashua (2008) solve a constrained convex optimization over an entropy error measure and solve it by the Sinkhorn iterative approximation, under a probabilistic view of the graph matching problem.

Spectral matching (SM) Leordeanu and Hebert (2005a) used pairwise constraints with the general formulation of graph matching and solved as a Quadratic Assignment Problem (QAP). By using the principal eigenvector of the assignment matrix they find a solution for the assignment, then iteratively removing outliers while and imposing the mapping constraints (one-to-one or one-to-many).

Spectral matching with affine constraints (SMAC) By adding affine constraints to the SM formulation Cour, Srinivasan and Shi (2007) extended SM improving scalability and also solved it as an eigenvalue problem.

Integer projected fixed point method (IPFP) Leordeanu, Hebert and Sukthankar (2009) proposed an integer projection algorithm with climbing and convergence properties to optimize the objective function in an integer domain. The algorithm is a sequence of linear assignment problems that finds the following binary solution from a previous one, and have two different versions, one initialized with uniform correspondence (IPFP-U) and other with spectral matching (IPFP-S).

Re-weighted random walk matching (RRWM) Cho, Lee and Lee (2010a) solved the graph matching problem based on simulating random walks with reweighting jumps widely used in Web ranking methods enforcing the matching constraints on the association graph.

Factorized graph matching (FGM) In their research Zhou and De la Torre (2016a) shown that the affinity matrix can be factorized to solve the graph matching problem, this reduces the computational cost for the proposed method and also for other methods in the literature to use the same approach. The algorithm has two versions one for directed (FGM-D) and other for undirected graphs (FGM-U).

Table 3 summarizes the presented graph matching methods with details and code source for each of them.

Table 3 – Graph matching comparison methods summary.

Method	Solver	Reference	Implementation
GA	continuation method	(GOLD; RANGARAJAN, 1996a)	(COUR PRAVEEN SRINIVASAN, 2009)
PM	Sinkhorn iterative approximation	(ZASS; SHASHUA, 2008)	(ZHOU; De la Torre, 2016b)
SM	spectral + iterative	(LEORDEANU; HEBERT, 2005a)	(LEORDEANU, 2009)
SMAC	eigenvalue problem	(COUR; SRINIVASAN; SHI, 2007)	(COUR PRAVEEN SRINIVASAN, 2009)
IPFP	iterativ. solves linear assingm.	(LEORDEANU; HEBERT; SUKTHANKAR, 2009)	(LEORDEANU, 2009)
RRWM	random walks	(CHO; LEE; LEE, 2010a)	(ZHOU; De la Torre, 2016b)
FGM	concave-convex procedure + Frank–Wolfe algorithm	(ZHOU; De la Torre, 2016a)	(ZHOU; De la Torre, 2016b)

6.2 Registration methods

The effectiveness and accuracy of our framework is verified by a comprehensive set of comparisons against eight robust image registration methods:

- *SIFT-based* by [Lowe \(2004\)](#);
- *SURF-based* by [Bay, Tuytelaars and Gool \(2006\)](#);
- *GDB-ICP* by [Yang et al. \(2007b\)](#);
- *Harris-PIIFD* by [Chen et al. \(2010\)](#);
- *ED-DB-ICP* by [Tsai et al. \(2010b\)](#);
- the technique proposed by [Chen et al. \(2011a\)](#) (namely *RIR-BS* for simplicity);
- Active Testing Search for Robust Graph Matching (*ATS-RGM*) by [Serradell et al. \(2015\)](#);
- and the recently published work by [Hernandez-Matas, Zabulis and Argyros \(2017\)](#) (which we refer here as *H-M*).

SIFT and SURF-based In order to compare our method with a baseline approach when dealing with image processing, we defined a simple pipeline with SIFT and SURF to find and describe keypoints followed by brute force matching and then an reliable outlier detection method (WGTM ([IZADI; SAEEDI, 2012](#))) and transformations that we used in our method. This should be an interesting comparison since SIFT and SURF are very reliable and used in multiple contexts of the image processing field and may provide a solid comparison material.

Generalized dual-bootstrap iterative closest point (GDB-ICP) GDB-ICP [Yang et al. \(2007c\)](#) tries to register the images first by identifying from a small overlapping area (bootstrap), using SIFT to find an initial transformation, then iteratively growing the bootstrap region using an iterative closest point (ICP) algorithm. GDB-ICP is even used in a commercial application (DualAlign) to register retinal images, however, GDB-ICP is susceptible in the presence of anatomical changes and also is affected by noise.

Harris-partial intensity invariant feature descriptor (Harris-PIIFD) Harris-PIIFD framework proposed by [Chen et al. \(2010\)](#) uses the Harris detector to find corner points as keypoints and from these are extracted the PIIFD descriptor which is invariant to rotation, partially to intensity, to perspective and affine transformations. Then, these points are matched, refined and adaptive transformation is used to register the image pairs.

Edge-driven dual-bootstrap iterative closest point (ED-DB-ICP) [Tsai et al. \(2010a\)](#) introduced a modification of GDB-ICP called ED-ED-ICP to register fundus images including

multimodal fluorescein angiograms (FA). In order to do that they used the gradient magnitude of the images to extract the SIFT keypoints and also added the edge points to the descriptors, hence creating a better more significant features in the context of multimodal images but also being able to register fundus images in general.

Retinal image registration using bifurcation structures (RIR-BS) [Chen *et al.* \(2011b\)](#) used the topological structure of the blood vessels to guide their matching in RIR-BS, where first the vessels are segmented and the bifurcations are found, then descriptors are made using the normalized branching angle and length, those features have the properties of invariance to translation, rotation, scale and according to the authors even a modest invariance to distortion. The method is very simple but efficient and can be easily applied to any type of fundus images.

Active testing search for robust graph matching (ATS-RGM) ATS-RGM was proposed by [Serradell *et al.* \(2015\)](#) and it's composed by two steps, a coarse alignment, and a finer alignment. The coarse alignment is done by graph matching, where first a random set of correspondences is determined and that is used to find the scale and orientation, and then a Gaussian Process is used to recursively refine the process. Next in the fine alignment is done using the geodesic distances among edges of the graphs are conserved constrained by the graph topology and the fine assignment is solved using the Hungarian algorithm.

Hernandez-Matas particle swarm optimization method (H-M) [Hernandez-Matas, Zabulis and Argyros \(2017\)](#) defined a method to register fundus images that initialize the problem using RANSAC and using a spherical eye model to estimate the camera pose, next relies on three independent Particle Swarm Optimizations (PSO) to search for the multidimensional pose and shape space. Keypoints are found and described by SIFT and the bifurcation structures and the best PSO result is chosen and applied to the image considering the best optimization score obtained. Additionally, to enable the use of GDB-ICP and ED-DB-ICP for high-resolution images, we operate them at coarser scales of the fundus photos. Also the best parameters are set accordingly the author's suggestions in their works or in their own implementations. [Table 4](#) was made to summarize these methods objectives and also provide the source of codes of each method.

Table 4 – Image registration methods.

Method	Type of registration	Reference	Implementation
SIFT-based	General images	(LOWE, 2004)	(VEDALDI; FULKERSON, 2010)
SURF-based	General images	(BAY; TUYTELAARS; GOOL, 2006)	(MATHWORKS, 2017)
GDB-ICP	General and fundus images	(YANG <i>et al.</i> , 2007b)	(YANG <i>et al.</i> , 2007a)
Harris-PIIFD	Multimodal fundus images	(CHEN <i>et al.</i> , 2010)	*
ED-DB-ICP	Multimodal fundus images	(TSAI <i>et al.</i> , 2010b)	*
RIR-BS	Fundus images	(CHEN <i>et al.</i> , 2011a)	(CHEN, 2009)
ATS-RGM	Fundus images and 3D Neuronal stacks	(SERRADELL <i>et al.</i> , 2015)	(SERRADELL <i>et al.</i> , 2015)
H-M'17	Fundus images	(HERNANDEZ-MATAS; ZABULIS; ARGYROS, 2017)	(HERNANDEZ-MATAS <i>et al.</i> , 2017a)

* Implementation provided by the authors.

6.3 Metrics

In order to quantitatively assess the quality of the registration results obtained by many methods, we employ a registration error measure which requires the corresponding ground truth points for each image pairs.

Mean Registration Error (MRE) (HERNANDEZ-MATAS *et al.*, 2017c) computes the error between the available prescribed points (ground truth) in image R and the corresponding ones in the estimated transformed image $T(I)$. More precisely, given N reference points $\mathbf{Y}_I = \{\mathbf{y}_I^i \in \mathbb{R}^2, i = 1, \dots, N\}$ and $\mathbf{Y}_R = \{\mathbf{y}_R^i \in \mathbb{R}^2, i = 1, \dots, N\}$ initially marked as matching points in I and R , we obtain the MRE measurement as follows:

$$MRE(\mathbf{Y}_I, \mathbf{Y}_R, T) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{y}_R^i - T(\mathbf{y}_I^i)\|_2, \quad (6.1)$$

where $\|\cdot\|_2$ denotes the Euclidean norm. Since MRE is an error measure, values closer to 0 represents more accurate results.

MRE's area under the curve (AUC) (HERNANDEZ-MATAS; ZABULIS; ARGYROS, 2017) provides an efficient manner to summarize the registration success for a given dataset. The MRE is a reliable measure for our experiments, but simply computing the mean and standard deviation of the set of registered image pairs may originate in biased results. Any of the methods presented may fail to register some image pairs, giving no output as a result or resulting in images projected far from the starting pose (outlier) these cases may skew standard

statistics. An alternative is plotting a continuous curve of the graph of the *success rate* (the ration between the number of successfully registered pairs by the total number of pairs) as a function of target accuracy, where the abscissa represents the value of an error threshold and the ordinate the success rate with MRE lower than the corresponding threshold. Next, the AUC is computed by the trapezoidal rule and scaled to the range $[0, 100]$. The higher scores indicate better results.

6.4 Datasets

We run all the evaluated techniques in three fundus image datasets with retinal pairs differing from each other in terms of geometric transformations as well as rendering aspects such as noise occlusion and low image quality. In fact, translation, rotation and scale variability are generated by involuntary movements of the patient eyes during the acquisition process of the fundus images. So, among the noisy evaluated images, some of them have low sharpness caused by lack of focus while a full set of image pairs present anatomical changes, a behavior typically

found in patients with some degenerative disease. Therefore, all the datasets aim at reflecting practical situations as described above.

Dataset 1 contains 18 pairs of high-resolution images available in a public repository by [Kohler et al. \(2013\)](#). This dataset presents twin-like images extracted from the same patient so that the images vary by an RST (Rotation-Scale-Translation) and a large amount of noise. Moreover, each retinal pair is composed of a good quality image with its counterpart impaired by specular noise or decreased sharpness.

Dataset 2 is formed by 85 image pairs collected by an ophthalmologist during his preventive exams for different patients. It covers real situations where undesirable artifacts were created during the capture of the digital fundus images. In fact, this dataset reveals an interesting overview of real cases faced by ophthalmologists and other eye specialists in their daily routine, so any registration algorithm useful in practice should be able to cope with cases gathered in this dataset.

FIRE dataset comprises 134 pairs of retinal images made publicly available by [Hernandez-Matas et al. \(2017a\)](#). FIRE is composed of three major classes of images: A, S and P. Category A contains images that present anatomical changes caused by degenerative eye conditions while Category S holds the opposite situation. Finally, the Category P covers the case where the pairs are recorded as mosaic representations of the eyes. Mosaics are sets of images used to provide a combined view of the entire retina by applying image stitching of different fundus images from same patient, hence, this images have low overlap. FIRE also provides ground truth and their corresponding points annotated by a specialist for each pair of images from the dataset.

In our experiments, regarding the registration error ([section 6.3](#)), were manually marked ground truth corresponding points in Datasets 1 and 2 following the same pattern as provided by the FIRE dataset.

Finally, in order to compute the registration error from a given set of ground truth points, we follow [Hernandez-Matas et al. \(2017a\)](#) towards delivering ten reference points and their correspondences for each pair of images for both Datasets 1 and 2. Additionally, the matchings have been carefully analyzed, corrected (when required) and finally approved by two experienced ophthalmologists (20 and 30 years of experience, respectively) to ensure high accuracy and validity of the correspondences.

6.5 Pipeline steps evaluation: a quantitative assessment

To better evaluate our contributions and to justify our choices, we compare our pipeline in separate modules starting with the feature set used in our cost function.

6.5.1 Feature selection

In order to define the best combination of features to use we tested all combination of the six features and computed the MRE for the three sets in Dataset 3.

Table 5 presents the error measured for the six best combination sets. First, we present the number of images registered with an error less than a threshold error, where 25 is the maximum error to consider the image registered according to Hernandez-Matas *et al.* (2017a). Next, we show the weighted average (by the vector [5, 4, 3, 2, 1]) since images registered with lower errors are preferred, lastly we show the AUC of the thresholded error (from 0 to 25 with an increment of 1) test to summarize our results. One can observe that feature set one and three yielded the best results but set three remains consistent over the three image subsets including A which is the one with patients with diseases which is extremely relevant in this context. For the final version of VOTUS, the feature set adopted was: gray patches, salient-based vector and the Gabor filter patches.

Dataset 3 contains a variety of images with noise, diseases, and normal patients and also mosaic cases, making it a reliable testing set. With the same mindset of machine learning testing set (to avoid overfit by selecting the best result in a single set), we choose the best set in this dataset and, as our results will show in subsection 6.5.5, the accuracy is still maintained in the other datasets, our choice should be a good general set for any fundus image registration dataset. Testing in a single dataset avoids a choice that is an “overfit” for all our datasets and hence not a good general choice.

Table 5 – Number of images registered varying the features composing the cost matrix, having the best case in bold and the second best underlined

Thresholds	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
Error<5	97	<u>100</u>	103	98	96	<u>100</u>
Error<10	121	118	121	<u>119</u>	118	115
Error<15	127	122	<u>124</u>	122	<u>124</u>	119
Error<20	129	125	<u>127</u>	124	126	123
Error<25	130	125	<u>128</u>	125	127	125
Weighted Avg.	<u>113.02</u>	112.08	114.82	111.31	110.90	110.76
Subset A AUC	0.65571	0.69857	<u>0.68143</u>	0.65286	<u>0.68143</u>	0.66143
Subset P AUC	0.70286	0.63714	<u>0.67184</u>	0.63878	0.63469	0.61429
Subset S AUC	<u>0.93268</u>	0.93042	0.93437	0.93437	0.93155	<u>0.93268</u>

SIFT
 BRISK
 Node position
 Gray patches
 Saliency vector
 Gabor patches

6.5.2 Graph matching comparison

First, VOTUS is analyzed in relation to the state-of-the-art graph matching methods, when they are applied in the context fundus image registration. The implementation of these methods are provided by Zhou and De la Torre (2016a) and it includes: *Graduated Assignment* (GA) (GOLD; RANGARAJAN, 1996b), *Spectral Matching with Affine Constraints* (SMAC) (LEORDEANU; HEBERT, 2005b), *Re-weighted Random Walk Matching* (RRWM) (CHO; LEE; LEE, 2010b) and *Factorized Graph Matching* (FGM) (ZHOU; De la Torre, 2016a). In this experiment, we only modify the graph matching stage in our pipeline (section 4.3), replacing our OT-based graph matching (OT-GM) by other methods. The same graphs were used as input for all methods, and their parameters were set as default. The result is shown in Table 6, we can observe that our strategy is superior in all datasets, specially on images from FIRE dataset due to the presence of noise from the structural changes caused by eye diseases (Category A) and the large number of graphs with topological changes (Category P).

6.5.3 Outlier removal comparison

Next, we analyze our choice for the outlier removal procedure, comparing the improved DeSAC filtering (section 4.7) against another robust method called *Weighted Graph Transformation Matching* (WGTM) (IZADI; SAEEDI, 2012). The key idea of WGTM relies on the application of an angular distance between the K -nearest-neighbor (K -NN) features. In several cases, WGTM outperforms the well-known RANSAC due to the fact that the angular distance is less sensitive to the noise caused by lens distortion. In particular, we use the best parameters found for WGTM ($K = 20$ and stopping threshold $\varepsilon = 10^{-3}$). Table 7 shows that our improved DeSAC achieved the best results and also has the benefit of being deterministic. This step is essential to achieve the best accuracy, notice that DeSAC improves significantly the results on images with the presence of noise (Category A of FIRE).

Table 6 – Comparative results using AUC for each graph matching method in our pipeline. Best results in bold.

Methods	FIRE Dataset			Datasets	
	A	P	S	1	2
GA	63.42	48.85	92.08	78.66	92.91
PM	0.0	0.0	11.83	01.44	18.47
SM	41.42	42.57	89.09	76.44	85.95
SMAC	64.28	58.08	92.25	78.44	92.89
IPFP-U	47.0	50.40	90.73	77.55	89.81
IPFP-S	47.28	35.55	88.19	63.77	82.18
RRWM	63.71	54.00	92.42	78.00	92.56
FGM	60.85	54.53	92.59	78.44	93.24
our OT-GM	68.14	67.18	93.43	78.66	94.56

Table 7 – Comparative results using AUC for different strategies of outlier removal in our pipeline. Best results in bold.

Methods	FIRE Dataset			Datasets	
	A	P	S	1	2
WGTM	52.42	63.63	93.04	78.00	94.47
our DeSAC	68.14	67.18	93.43	78.66	94.56

Table 8 – Statistics of the registration methods in their both standard and GC-optimized versions when assessed by the MRE measure.

Thresholds	SIFT-based		SURF-based		GDB-ICP		ED-DB-ICP		RIRBS	
	Standard	GC	Standard	GC	Standard	GC	Standard	GC	Standard	GC
Mean	6.5689	4.8962	6.1012	4.8187	10.301	5.9499	8.3293	4.3363	6.9646	4.4631
Median	4.5484	3.1705	3.8692	3.0966	8.7278	4.0437	6.5854	2.2714	5.1057	3.4790
Std	4.9817	4.6785	5.9422	4.6242	5.8711	4.8703	6.0571	5.1967	5.4473	3.0553

6.5.4 Best fit transformation results

To validate the performance of our approach towards optimizing registration algorithms that are typically found in the field of fundus photography, we conduct a comprehensive experimental evaluation taking five commonly used methods and a representative public benchmark.

More precisely, we make use of the FIRE dataset (HERNANDEZ-MATAS *et al.*, 2017b) which contains 134 image pairs reflecting a considerable number of eye impairments and other aggravating factors such as the presence of noise, eye’s involuntary movements, and visual anatomical changes. For each image pair, its available a set of 10 corresponding points marked by a medical expert which is taken as “ground truth” to calculate the MRE, given in pixels. We then assess the accuracies reached by the following registration methods with/without our boosting scheme: SURF (CATTIN *et al.*, 2006) and SIFT (M; MOSES, 2014) tuned with RANSAC (DERPANIS, 2010) for outliers removal, GDB-ICP (YANG *et al.*, 2007c), ED-DB-ICP (TSAI *et al.*, 2010a) and RIR-BS (CHEN *et al.*, 2011b). Finally, following Hernandez-Matas *et al.* (2017b), we also consider that an image pair is indeed registered whether the MRE score is at least less than 25.

Notice that this is the only evaluation where we can use statistical measures and boxplots, because the comparison methods have the same number of registered images (since each method is compared with itself), hence the statistics are not influenced by non-registered images.

We start our experimental analysis by extracting basic statistics like *mean*, *median* and *standard deviation* for the MRE over all the evaluated methods in their usual and GC-optimized versions, i.e., when the methods are combined with our GC-driven scheme for optimal mapping selection. From the statistics listed in Table 8, one can verify that GC leads to a significant improvement in the registration accuracy, attesting its good performance in all the measurements. For example, the error computed for the conventional algorithms GDB-ICP and ED-DB-ICP are reduced to approximately half when they are equipped with our boosting scheme.

Table 9 – Number of registered images for different parameters of the MRE for both standard and GC-driven registration methods.

Thresholds	SIFT-based		SURF-based		GDB-ICP		ED-DB-ICP		RIRBS	
	Standard	GC	Standard	GC	Standard	GC	Standard	GC	Standard	GC
Error<5	61	81	49	52	42	85	22	64	35	52
Error<10	91	99	64	67	67	97	60	94	57	67
Error<15	99	104	68	72	87	104	78	103	67	77
Error<20	110	113	70	75	96	113	96	114	71	79
Error<25	112	115	76	77	102	119	102	121	74	82

Figure 16 – MRE for the registration methods when setting different geometric transformation models.

A deeper and more prescriptive evaluation concerning the registrations produced by the techniques can be visualized in Table 9. Again, our GC-driven scheme significantly increases the accuracy performance for all the methods, performing a larger number of registrations even in more drastic cases such as the ones posed by low MRE thresholds. Another interesting aspect to be observed is that in certain circumstances, where a conventional method registers much more images than others, it is possible to reverse this scenario just by enabling the GC scheme into the second group of techniques (e.g., see SIFT-based vs. ED-DB-ICP).

In Figure 16, we assess the accuracy of the registration methods separately, i.e., by setting different transformations. More specifically, the methods are evaluated according to the following

Figure 17 – Number of times that a certain geometric transformation model is chosen as optimal by our CG scheme and the MRE.

set of geometric models: *similarity*, *affine*, *homography*, *quadratic*, and, finally, by means of our GC-based decision rule. By checking the boxplots in Figure 16, one can see that the registrations delivered by our scheme have the lowest MRE values, overall. The SURF method didn't return the minimum number of collinear points demanded to properly establish a quadratic model. We also compute in Figure 17 the frequency at which each geometric transformation is chosen as optimal by our GC scheme. The best choice measured by the MRE is also provided as “ground truth”. The choices made by our approach follow the ones achieved by the MRE, being very stable regarding the MRE distribution for all the mappings. Another observed behavior for both GC and MRE metrics is that there is no absolute transformation to any registration method: they assume different models as the image pairs vary.

Next, we discuss how the designed GC scheme behaves in terms of *consistency*. Figure 18 displays the histogram of the geometric mappings chosen by our approach. Note that, as the best choice, GC highly maximizes the MRE, followed by the second-best choice, and so on. In other words, the registration methods driven by the GC scheme reach a good level of consistency, where the probability of choosing an inaccurate transformation falls highly according to the expected, indicating that the presented method is stable and can be trusted when estimating unseen samples.

We conclude this section showing in Figure 19 a qualitative set of results for a pair of retinal images. For each registration method, we consider two scenarios: the usual mappings, adopted by the current methods, versus the best competitor, as properly computed by our GC scheme. As shown by the visual and quantitative results measured by the MRE and CG, the proposed approach archives a satisfactory performance in terms of registration accuracy, varying from incremental to substantial improvements (e.g., see the SIFT and GDB-ICP examples).

Figure 18 – Number of transformations selected by the GC from best (Rank 1) to worst (Rank 4). Clearly, the histograms show that GC tends to determine the best transformations for each method.

6.5.5 Image registration comparison

We compare VOTUS against eight robust image registration methods on all datasets (section 6.4), from well-established keypoint-based methods SIFT-based (LOWE, 2004), SURF-based (BAY; TUYTELAARS; GOOL, 2006), and GDB-ICP (YANG *et al.*, 2007b), to techniques designed to formally accomplish fundus image alignment, i.e., Harris-PIIFD (CHEN *et al.*, 2010), ED-DB-ICP (TSAI *et al.*, 2010b) and H-M'17 (HERNANDEZ-MATAS; ZABULIS; ARGYROS, 2016) and also compare with two fundus registration techniques that use graphs RIR-BS (CHEN *et al.*, 2011a) and ATS-RGM (SERRADELL *et al.*, 2015).

In order to prove that our method is effective and reliable, we compare with a standard image registration approach which uses SIFT or SURF for keypoints discovery and description united with a robust outlier removal technique (WGTM) (IZADI; SAEEDI, 2012). Also, after the computation of RANSAC, the registration is effectively reached also by the scheme presented by Motta, Casaca and Paiva (2018).

The implementation of most of these methods are available online or provided by the authors. For each method, we choose the best parameters following the recommendations stated by the authors as found in the source code. These methods and their implementations were compiled in Table 4.

6.5.6 Qualitative analysis

We start our comparative analysis showing how visual inspection can be accomplished when viewing the registration results. Figure 20 illustrates the process of inspecting the matching between the fundus images by counting the amount of overlap pixel remaining in the registered

Figure 19 – Qualitative comparison for a given image pair when assessed by all the five evaluated methods.

images. More specifically, we decompose into RGB channels the binary vessels of the fundus images B_I and B_R so that the green component $(0, 1, 0)$ is taken from B_I while its counterpart, the magenta color $(1, 0, 1)$, is designated to the binary image B_R . These cropped components are then overlaid so that the resulting image will have white pixels $(1, 1, 1) = (0, 1, 0) + (1, 0, 1)$ denoting a vessel match, and green/magenta colors, otherwise. Therefore, higher amounts of white pixels in the composition (or the absence of green and magenta) indicate better results.

Figure 20 – Color labeling-based inspection: the composition of two binary vessel images B_I (green) and B_R (magenta) so that complementary color channels are used to drive the visual analysis. A more substantial amount of white pixels means better registration outcomes.

The flexibility and effectiveness of VOTUS to cope with different kind of images are

illustrated in Figure 21. Notice from the results that a significant number of vessels was accurately assigned by our approach, attesting the accuracy of the method for the cases where the images differ each other as to color variation, the eminence of artifacts and the presence of geometric deformations. Even for the more severe situation, shown in the second example, our technique could align most vessel structures in the images. Considering the visual inspection of the outcomes for the particular cases of diseases, pathogenic disorders or even unexpected structural changes on the retinal blood vessels, VOTUS provides an interesting perception of the eye fundus alignment as two rightmost cases. For instance, the number of green veins located in the northwest region in the last sequence of eye images could be successfully used to detect unexpected vessels changes as a portion of veins has disappeared from the secondary validated image.

As stated by Ghassabi et al. (GHASSABI; SHANBEHZADEH; MOHAMMADZADEH, 2016), an essential aspect of registration techniques is their ability in handling high-resolution images. For instance, by examining these images in such a way that the resulting alignment is inspected at its original scale, one could track the progress of eye diseases such as *diabetic retinopathy*¹. Other benefits are the use of compatible and more advanced technologies to store and retrieve the results as well as powerful visual exploratory systems to accurately audit the images. Figure 22 depicts qualitative comparisons of VOTUS against five methods over four high-resolution images taken from Dataset 1.

As illustrated, GDB-ICP, ED-DB-ICP, and SURF+WGTM produce acceptable results, but they carry some miss-alignments (noticed by magenta and green pixels rather than white), and most methods (except VOTUS) generate misfitted alignments at least on the first-row image for which the original image had some focus noise. In contrast, VOTUS generates more consistent and pleasant results in all cases, mainly regarding the quality of matching refinement as shown by a large amount of white color in the resulting montages.

In order to quantify the amount of white pixel on registered images, we use the GC measure (Equation 5.4) to obtain this estimative. Figure 23 shows the boxplots of the GC scores (high values are best) in Dataset 1. We can observe by the boxplots that VOTUS outperforms all methods. However, although GC provides a reasonable estimative for the quality of the registration, the main drawback is that GC is not an entirely reliable measure for the registration success, as seen by the low gain reported by H-M'17 resulted from interpolation artifacts. In the next section, we will compare the registration methods using reliable measures, such as MRE and AUC (see subsection 6.5.7).

6.5.7 Quantitative analysis

Table 10 depicts a set of comparisons of our approach versus a class of challenging registration techniques (see Table 4) mainly devised to align fundus images on all three datasets

¹ Eye disease that occurs as a complication of diabetes that can lead to blindness (GROUP et al., 2004).

Figure 21 – Visual results generated by VOTUS when applied to arbitrary images with different losses of quality from Dataset 1 (first two columns) while the remaining ones ticked with anatomical changes, which have been collected from FIRE.

considered. For the non-deterministic methods, such as GDB-ICP, H-M'17, ATS-RGM, we run the method three times and take the average of MRE.

As can be seen, VOTUS leads to higher AUC scores, demonstrating its accuracy and robustness in a substantial number of cases. It is more evidenced in Dataset 1 which contains noisy images and also in Category P of FIRE dataset which includes mosaic images, where the significant difference of quality occurs due to our OT-based graph matching combined with our improved DeSAC filtering. Looking at the results of Dataset 2 and Category A of the FIRE dataset, our approach also outperforms the other methods. Indeed, our approach is effective when dealing with anatomical changes in the eye fundus, thus attending to medical protocols for real ophthalmic applications. The excellent performance of VOTUS is also endorsed in image pairs

Table 10 – Comparative results of VOTUS against other image registration methods using AUC. The best results and the second best are shown in bold/red and italic/blue, respectively.

Methods	FIRE Dataset			Datasets	
	A	P	S	1	2
VOTUS	68.14	67.18	93.43	78.66	94.56
H-M'17	<i>56.86</i>	50.29	<i>93.38</i>	<i>68.56</i>	<i>94.09</i>
GDB-ICP	29.57	30.24	79.77	23.00	89.55
Harris-PIIFD	42.71	07.55	88.02	03.00	58.14
ED-DB-ICP	49.71	44.12	60.39	43.00	62.80
SIFT+WGTm	40.71	<i>54.40</i>	83.69	67.33	90.63
SURF+WGTm	06.85	06.08	83.54	52.00	72.23
RIR-BS	12.42	04.93	77.21	49.55	68.00
ATS-RGM	02.14	00.00	04.81	00.00	25.17

of Dataset 2 and Category S (pairs without anatomical changes), H-M'17 and VOTUS are the methods that deliver the best performance even when H-M'17 almost matches with VOTUS.

However, one must consider that H-M'17 is not a deterministic method. Indeed, H-M'17 uses a combination of non-deterministic methods, such as particle swarm optimizations and RANSAC, applied many times before getting the best score regarding its minimization function score. For instance due to this randomness, given an image pair, we run three times the H-M'17 method in this pair and obtain an average MRE score of 21.03 with a standard deviation of 8.01, the results are shown in Figure 24. In some cases, H-M'17 produces spurious results (Figure 24d), failing in registration (with MRE greater than 25) and producing interpolation artifacts, the curved vertical lines wrongly detected as vessels. While VOTUS is a fully automated and stable algorithm that accurately performs the registration even under several variate conditions of noise, essential traits desirable in many practical applications, providing a steady MRE score of 7.65 (Figure 24a).

Figure 25 presents the boxplots of the MRE for VOTUS as well as for H-M'17, this confirms that VOTUS is more stable due to the low variability of the MRE scores depicted by the interquartile range and still avoiding the large uncertainties (as illustrated by the upper whiskers) associated with the registration in some image pairs.

6.5.8 Statistical hypothesis testing

In order to verify if the results afforded by VOTUS are statistically more significant than the results provided by the other methods, we perform the Student's t-test in these results. The p-value is the level of relevance computed from the Student's t-distribution, and low p -values ($p < 0.05$) indicate that the results are statistically more accurate compared with the rest of the methods. Our null hypothesis is that $H_0 : \bar{V}_\varepsilon \leq \bar{O}_\varepsilon$, where $\bar{V}_\varepsilon$ is the arithmetic mean of the number of images registered by VOTUS under a threshold error ε varying from 0 to 25 (with increment of 0.25) and $\bar{O}_\varepsilon$ is the respective mean for the other comparison methods. In

particular, we reject H_0 and state that our method statistically register more images on average than any other method for a given dataset if the corresponding p -value is low. Table 11 shows the p -values computed per dataset and the last column represents the number of times that H_0 was rejected. Notice that VOTUS is statistically superior in at least three datasets in comparison to each method. In another hand, we also check the otherwise and no other method proved to be statistically more significant than VOTUS in any dataset.

Table 11 – The p -values of the registration results by the other methods in comparison with VOTUS.

Methods	FIRE Dataset			Datasets		#Rej. H_0
	A	P	S	1	2	
H-M'17	0.005	0.000	0.508	0.016	0.436	3
GDB-ICP	0.000	0.000	0.000	0.000	0.049	5
Harris-PIIFD	0.000	0.000	0.067	0.000	0.000	4
ED-DB-ICP	0.000	0.000	0.000	0.000	0.000	5
SIFT+WGTm	0.000	0.000	0.002	0.013	0.099	4
SURF+WGTm	0.000	0.000	0.002	0.000	0.000	5
RIR-BS	0.000	0.000	0.000	0.000	0.000	5
ATS-RGM	0.000	0.000	0.000	0.000	0.000	5

6.6 Computational aspects

We implemented our method in MATLAB R2015a and all experiments have been performed on a laptop equipped with Intel Core i7 processor with 8Gb of RAM.

Table 12 shows the averaging computational timings per image pair registration in Dataset 2. Our approach is considerably faster than H-M'17 and slower than other methods (in an ideal context for which one single RANSAC step is executed). Notice that, our method is 25 times faster than the competitor with the best accuracy, H-M'17, using CPU and still almost two times faster than the time reported by the authors using GPGPU, although our method could get a lot faster by using parallelism as well. The traditional methods (GDB-ICP, ED-DB-ICP, and Harris-PIIFD) are faster than our approach, but does not provide accurate results under real-world conditions (see subsection 6.5.7). The accuracy of VOTUS comes with a price of solving the optimizations required. However, the average time presented is feasible even being executed along a physician consultation, attending the ophthalmologic protocol.

Table 12 – Computational times (in seconds).

Methods	VOTUS	H-M'17	ED-DB-ICP	Harris-PIIFD	GDB-ICP
Time	106	2674	44	13	19

Figure 22 – Qualitative comparisons from sample images collected from Dataset 1.

Figure 23 – Boxplots of the GC scores in Dataset 1 for the methods: VOTUS (■), H-M'17 (■), GDB-ICP (■), ED-DB-ICP (■), SIFT+WGTM (■), SURF+WGTM (■) and RIR-BS (■).

Figure 24 – MRE for a given image pair from Dataset 1.

(a) VOTUS, MRE = 7.65

(b) H-M'17, MRE = 12.16

(c) H-M'17, MRE = 23.16

(d) H-M'17, MRE = 27.75

Figure 25 – Boxplots of the MRE scores for VOTUS (■) and H-M'17 (■) in all datasets.

CONCLUSIONS

In this chapter we describe our study conclusions and highlight the limitations of VOTUS and provide our plans for future researches.

7.1 Discussion and limitations

Comparisons presented in [Chapter 6](#) demonstrate the accuracy and effectiveness of VOTUS in dealing with a wide variety of retinal fundus images on representative medical datasets. In fact, VOTUS can deal with color variations, generating pleasant results regarding many image quality aspects such as lack of focus, illumination distortion, and blood vessel motion — a particular and desirable characteristic not always present in most existing registration techniques. The alignments resulting from our approach turned out to be clean, well assembled, and easy to check. Regarding the mathematical facet of our method, the proposed OT scheme has proven to be quite robust and flexible, providing very accurate matchings while guaranteeing a unique solution due to the mathematical formulation.

Despite its robustness, versatility and solid mathematical foundation it relies on, there are two aspects must be discussed involving the input image pairs (see [Figure 26](#)).

First, the registration of the retinal images may be affected by the absence of a salient region to use as a reference, such as the ONH, which can affect the performance of the method. VOTUS uses this information as a feature to build the cost matrix, and without it, our method could fail. However, in some cases, our method had successful to register images even without the ONH information.

The second and most important aspect is related to the information of the blood vessels to generate the graph and if this information is missing due to an extreme lack of focus. This issue is also a limitation of our method, but in this case, the image may be useless to the physician visual inspection as well.

Figure 26 – Limitations of VOTUS: images pairs without the ONH as reference (top row) and with lack of focus (bottom row).

7.2 Conclusions and future work

In this dissertation, we introduced VOTUS, a new technique that relies on the optimal mass transport theory to tackle the problem of fundus image registration. The proposed framework holds good accuracy in terms of changes in geometry vessel structures, lack of focus, and the presence of specular noise. Moreover, the technique makes more noticeable the effects of the degeneration in the eye's vessels, which can be caused by a variety of diseases.

As we demonstrated, VOTUS outperforms other evaluated state-of-the-art approaches even when compared against very robust and traditional registration methods. All those properties render VOTUS a useful and compelling retinal registration technique to aid physicians and ophthalmologists in practical cases.

During this research, we also revisited the common choice for the geometric transformation used by most methods in the literature, and we discover that the belief that a quadratic transformation is the best choice, under the assumption that it generalizes best the shape of the eye, is not really true for a large number of cases. We made our contribution by proposing a simplistic scheme that can decide the best transformation for each image pair, and have show in

our results improvements for a considerable number of methods with insignificant increase in the computational effort.

Figure 27 – VOTUS extended to multimodal fundus images: binary vessels overlapping image (left) and the classical checkerboard representation (right).

As future work, we are currently adapting our method to address multimodal fundus images registration ([GHASSABI *et al.*, 2013](#); [WANG *et al.*, 2015](#); [LEE *et al.*, 2015](#)). Originally, VOTUS was not developed to deal with multimodal images, however, we achieve promising results with a slight change in our method. The minor change consists in including in our pipeline the gradient of the images instead of the raw images and [Figure 27](#) shows some preliminary

results involving patients with diabetic retinopathy (dataset provided by [Alipour, Rabbani and Akhlaghi \(2012\)](#)). Next, we will perform a comparative analysis with the state-of-the-art methods of multimodal image registration.

In other directions, we intend to extend our method to compute 3D stereoscopic reconstruction of fundus images, which is another underlying problem related to the context of diagnostic assistance. Another possibility is to adapt our core pipeline to general images and solve the image stitching problem used, among other tasks, to do panoramic photos.

7.3 Bibliographic production related to this dissertation

These papers were made and submitted during this doctoral course.

1. ([MOTTA; CASACA; PAIVA, 2018](#)) D. Motta, W. Casaca, and A. Paiva - *Fundus Image Transformation Revisited: Towards Determining More Accurate Registrations*
CBMS 2018 – IEEE International Symposium on Computer-Based Medical Systems.
2. D. Motta, W. Casaca, and A. Paiva - *Vessel Optimal Transport for Automated Alignment of Retinal Fundus Images*
Submitted to TIP - IEEE Transactions on Image Processing.

7.4 Other bibliographic production

1. ([MOTTA et al., 2018](#)) D. Motta, W. Casaca, and A. Paiva - *A user-friendly interactive framework for unsteady fluid flow segmentation and visualization*
J Vis 2018 – Springer Journal of Visualization.
2. ([GOMEZ-NIETO et al., 2016](#)) E. Gomez-Nieto, W. Casaca, D. Motta, I. Hartmann, G. Taubin, L.G. Nonato - *Dealing with Multiple Requirements in Geometric Arrangements*
TVCG 2015 – IEEE Transactions on Visualization and Computer Graphics.
3. ([CASACA et al., 2015](#)) W. Casaca, D. Motta, G. Taubin, L.G. Nonato - *A User-Friendly Interactive Image Inpainting Framework using Laplacian Coordinates*
ICIP 2015 – IEEE International Conference on Image Processing.
4. ([MOTTA et al., 2015b](#)) D. Motta, M. Oliveira, P. Pagliosa, L.G. Nonato and A. Paiva - *Exploratory segmentation of vector fields using multidimensional projection*
SIBGRAPI 2015 – IEEE Conference on Graphics, Patterns and Images.
5. ([MOTTA et al., 2015a](#)) D. Motta, L. Matos, A.C. Souza, R. Marcato, A. Paiva and L.A.V. Carvalho - *All-in-focus imaging technique used to improve 3d retinal fundus image*

reconstruction

SAC 2015 – ACM Symposium On Applied Computing.

6. (MOTTA *et al.*, 2014) D. Motta, A. Serillo, L. Matos, F.M.M. Yasuoka, V.S. Bagnato and L.A.V. Carvalho - *3D papillary image capturing by the stereo fundus camera system for clinical diagnosis on retina and optic nerve*
SPIE BIOS 2014 – International Society for Optics and Photonics Proceedings.
7. (NIETO; MOTTA; NONATO, 2014) E. Gomez-Nieto, D. Motta and L.G. Nonato - *Semantically Aware Dynamic Layouts*
SIBGRAPI 2014 – IEEE Conference on Graphics, Patterns and Images.

BIBLIOGRAPHY

ABRAMOFF, M. D.; NIEMEIJER, M.; SUTTORP-SCHULTEN, M. S.; VIERGEVER, M. A.; RUSSELL, S. R.; GINNEKEN, B. van. Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes. **Diabetes Care**, v. 33, n. 8, p. 64–84, 2008. Citation on page 23.

ADAL, K. M.; ENSING, R. M.; COUVERT, R.; ETTEN, P. van; MARTINEZ, J. P.; VERMEER, K. A.; VLIET, L. J. van. A hierarchical coarse-to-fine approach for fundus image registration. In: **Proc. WBIR**. [S.l.: s.n.], 2014. p. 93–102. Citation on page 36.

ALIPOUR, S. H. M.; RABBANI, H.; AKHLAGHI, M. R. Diabetic retinopathy grading by digital curvelet transform. **Computational and mathematical methods in medicine**, Hindawi, v. 2012, 2012. Citation on page 80.

BANKHEAD, P.; SCHOLFIELD, C. N.; MCGEOWN, J. G.; CURTIS, T. M. Fast retinal vessel detection and measurement using wavelets and edge location refinement. **PloS One**, v. 7, n. 3, p. e32435, 2012. Citation on page 40.

_____. Fast retinal vessel detection and measurement using wavelets and edge location refinement. **PloS One**, v. 7, n. 3, p. e32435, 2012. Citation on page 54.

BAY, H.; TUYTELAARS, T.; GOOL, L. V. SURF: Speeded up robust features. In: **Proc. ECCV**. [S.l.: s.n.], 2006. p. 404–417. Citations on pages 36, 59, 60, and 68.

CAN, A.; STEWART, C. V.; ROYSAM, B.; TANENBAUM, H. L. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. **IEEE T. Pattern Anal.**, v. 24, n. 3, p. 347–364, 2002. Citation on page 35.

CASACA, W.; MOTTA, D.; TAUBIN, G.; NONATO, L. G. A user-friendly interactive image inpainting framework using laplacian coordinates. In: IEEE. **Image Processing (ICIP), 2015 IEEE International Conference on**. [S.l.], 2015. p. 862–866. Citation on page 80.

CATTIN, P. C.; BAY, H.; GOOL, L. V.; SZEKELY, G. Retina mosaicing using local features. In: **International Conference on Medical Image Computing and Computer-Assisted Intervention**. [S.l.: s.n.], 2006. p. 185–192. Citations on pages 23 and 65.

CHEN, J.; TIAN, J.; LEE, N.; ZHENG, J.; SMITH, R. T.; LAINE, A. F. A partial intensity invariant feature descriptor for multimodal retinal image registration. **IEEE T. Bio-med. Eng.**, v. 57, n. 7, p. 1707–1718, 2010. Citations on pages 36, 59, 60, and 68.

CHEN, L. **Feature-based retinal image registration**. 2009. [Http://www.mathworks.com/matlabcentral/profile/authors/1522396-li-chen](http://www.mathworks.com/matlabcentral/profile/authors/1522396-li-chen). Citation on page 60.

CHEN, L.; LIAN, Y.; GUO, Y.; WANG, Y.; HATSUKAMI, T. S.; PIMENTEL, K.; BALU, N.; YUAN, C. A vascular image registration method based on network structure and circuit simulation. **BMC Bioinformatics**, v. 18, n. 1, p. 229, 2017. Citation on page 35.

CHEN, L.; XIANG, Y.; CHEN, Y.; ZHANG, X. Retinal image registration using bifurcation structures. In: **IEEE Image Proc.** [S.l.: s.n.], 2011. p. 2169–2172. Citations on pages 35, 41, 59, 60, and 68.

_____. Retinal image registration using bifurcation structures. In: **IEEE International Conference on Image Processing (ICIP)**. [S.l.: s.n.], 2011. p. 2169–2172. Citations on pages 60 and 65.

CHIAPPORI, P.-A.; MCCANN, R. J.; NESHEIM, L. P. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. **Econ. Theor.**, v. 42, n. 2, p. 317–354, 2010. Citation on page 30.

CHO, M.; LEE, J.; LEE, K. M. Reweighted random walks for graph matching. In: SPRINGER. **European conference on Computer vision**. [S.l.], 2010. p. 492–505. Citation on page 58.

_____. Reweighted random walks for graph matching. In: **ECCV**. [S.l.: s.n.], 2010. p. 492–505. Citation on page 64.

COUR PRAVEEN SRINIVASAN, J. S. T. **Graph Matching Toolbox in MATLAB**. 2009. [Http://www.timotheecour.com/software/graph_matching/graph_matching.html](http://www.timotheecour.com/software/graph_matching/graph_matching.html). Citation on page 58.

COUR, T.; SRINIVASAN, P.; SHI, J. Balanced graph matching. In: SCHÖLKOPF, B.; PLATT, J.; HOFFMAN, T. (Ed.). **Advances in Neural Information Processing Systems 19**. MIT Press, 2007. p. 313–320. Available: <<http://papers.nips.cc/paper/2960-balanced-graph-matching.pdf>>. Citation on page 58.

DENG, K.; TIAN, J.; ZHENG, J.; ZHANG, X.; DAI, X.; XU, M. Retinal fundus image registration via vascular structure graph matching. **Int. J. Biomed. Imaging**, v. 2010, p. 14:1–14:13, 2010. Citations on pages 22 and 35.

DERPANIS, K. G. **Overview of the Ransac Algorithm**. [S.l.], 2010. Citation on page 65.

DIGNE, J.; COHEN-STEINER, D.; ALLIEZ, P.; GOES, F. D.; DESBRUN, M. Feature-preserving surface reconstruction and simplification from defect-laden point sets. **Journal of mathematical imaging and vision**, Springer, v. 48, n. 2, p. 369–382, 2014. Citations on pages 33 and 34.

DOMINITZ, A.; TANNENBAUM, A. Texture mapping via optimal mass transport. **Visualization and Computer Graphics, IEEE Transactions on**, IEEE, v. 16, n. 3, p. 419–433, 2010. Citations on pages 24, 33, and 34.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. [S.l.]: John Wiley & Sons, 2012. Citation on page 48.

EL-GAMAL, F. E.-Z. A.; ELMOGY, M.; ATWAN, A. Current trends in medical image registration and fusion. **Egyptian Informatics Journal**, v. 17, n. 1, p. 99–124, 2016. Citation on page 23.

ELDEEB, S. M.; FAHMY, A. S. Error analysis of fundus image registration using quadratic model transformation. In: **Int. Biomedical Engineering Conference**. [S.l.: s.n.], 2014. p. 137–140. Citation on page 53.

FANG, B.; TANG, Y. Y. Elastic registration for retinal images based on reconstructed vascular trees. **IEEE T. Bio-med. Eng.**, v. 53, n. 6, p. 1183–1187, 2006. Citation on page [35](#).

FERRADANS, S.; PAPADAKIS, N.; RABIN, J.; PEYRÉ, G.; AUJOL, J.-F. Regularized discrete optimal transport. In: **Scale Space and Variational Methods in Computer Vision**. [S.l.]: Springer, 2013. p. 428–439. Citations on pages [24](#), [32](#), and [33](#).

FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. **Commun. ACM**, v. 24, n. 6, p. 381–395, 1981. Citation on page [37](#).

_____. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Communications of the ACM**, ACM, v. 24, n. 6, p. 381–395, 1981. Citation on page [47](#).

GHASSABI, Z.; SEDAGHAT, A.; SHANBEHZADEH, J.; FATEMIZADEH, E. An efficient approach for robust multimodal retinal image registration based on ur-sift features and piifd descriptors. **EURASIP J. Image Vide.**, v. 2013, p. 25, 2013. Citations on pages [36](#) and [79](#).

GHASSABI, Z.; SHANBEHZADEH, J.; MOHAMMADZADEH, A. A structure-based region detector for high-resolution retinal fundus image registration. **Biomed. Signal Process. Control**, v. 23, p. 52–61, 2016. Citations on pages [35](#) and [70](#).

GOES, F. D.; COHEN-STEINER, D.; ALLIEZ, P.; DESBRUN, M. An optimal transport approach to robust reconstruction and simplification of 2d shapes. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2011. v. 30, n. 5, p. 1593–1602. Citations on pages [24](#), [33](#), and [34](#).

GOES, F. de; BREEDEN, K.; OSTROMOUKHOV, V.; DESBRUN, M. Blue noise through optimal transport. **ACM Transactions on Graphics (TOG)**, ACM, v. 31, n. 6, p. 171, 2012. Citations on pages [24](#), [32](#), and [33](#).

GOLD, S.; RANGARAJAN, A. A graduated assignment algorithm for graph matching. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 18, n. 4, p. 377–388, 1996. Citations on pages [57](#) and [58](#).

_____. A graduated assignment algorithm for graph matching. **IEEE PAMI**, v. 18, n. 4, p. 377–388, 1996. Citation on page [64](#).

GOMEZ-NIETO, E.; CASACA, W.; MOTTA, D.; HARTMANN, I.; TAUBIN, G.; NONATO, L. G. Dealing with multiple requirements in geometric arrangements. **IEEE transactions on visualization and computer graphics**, IEEE, v. 22, n. 3, p. 1223–1235, 2016. Citation on page [80](#).

GRAY, R. M. **Entropy and Information Theory**. [S.l.]: Springer, 2011. Citation on page [53](#).

GROUP, E. D. P. R. *et al.* The prevalence of diabetic retinopathy among adults in the united states. **Archives of ophthalmology**, Am. Med. Assoc., v. 122, n. 4, p. 552, 2004. Citation on page [70](#).

HAKER, S.; ZHU, L.; TANNENBAUM, A.; ANGENENT, S. Optimal mass transport for registration and warping. **International Journal of Computer Vision**, Springer, v. 60, n. 3, p. 225–240, 2004. Citations on pages [24](#), [31](#), [32](#), and [33](#).

_____. Optimal mass transport for registration and warping. **Int. J. Comput. Vision**, v. 60, n. 3, p. 225–240, 2004. Citation on page 30.

HARALICK, R. M.; SHAPIRO, L. G. **Computer and Robot Vision**. 2nd. ed. [S.l.]: Prentice Hall, 2002. ISBN 0201569434. Citation on page 40.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: **Proc. of the 4th Alvey Vision Conference**. [S.l.: s.n.], 1988. p. 147–151. Citation on page 36.

HARTLEY, R.; ZISSERMAN, A. **Multiple view geometry in computer vision**. [S.l.]: Cambridge university press, 2003. Citation on page 48.

HERNANDEZ-MATAS, C.; ZABULIS, X.; ARGYROS, A. A. Retinal image registration through simultaneous camera pose and eye shape estimation. In: **IEEE EMBC**. [S.l.: s.n.], 2016. p. 3247–3251. Citations on pages 36 and 68.

_____. An experimental evaluation of the accuracy of keypoints-based retinal image registration. In: **IEEE EMBC**. [S.l.: s.n.], 2017. (to appear). Citations on pages 37, 59, 60, and 61.

HERNANDEZ-MATAS, C.; ZABULIS, X.; TRIANTAFYLLOU, A.; ANYFANTI, P.; DOUMA, S.; ARGYROS, A. A. FIRE: Fundus Image Registration Dataset. **Journal for Modeling in Ophthalmology**, v. 1, n. 4, p. 16–28, 2017. Available: <<http://www.ics.forth.gr/cvrl/fire/>>. Citations on pages 60, 62, and 63.

_____. Fire: Fundus image registration dataset. **Journal for Modeling in Ophthalmology**, v. 1, n. 4, p. 16–28, 2017. Citation on page 65.

HERNANDEZ-MATAS, C.; ZABULIS, X.; TRIANTAFYLLOU, A.; ANYFANTI, P.; ARGYROS, A. A. Retinal image registration under the assumption of a spherical eye. **Comput. Med. Imag. Grap.**, v. 55, p. 95–105, 2017. Citations on pages 36 and 61.

HOOVER, A.; GOLDBAUM, M. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. **IEEE TMI**, v. 22, n. 8, p. 951–958, 2003. Citation on page 47.

IZADI, M.; SAEEDI, P. Robust weighted graph transformation matching for rigid and nonrigid image registration. **IEEE T. Image Proces.**, v. 21, n. 10, p. 4369–4382, Oct 2012. Citations on pages 48, 59, 64, and 68.

KAMARAINEN, J.-K.; KYRKI, V.; KÄLVIÄINEN, H. Invariance properties of gabor filter-based features-overview and applications. **IEEE T. Image Proces.**, v. 15, n. 5, p. 1088–1099, 2006. Citation on page 46.

KARALI, E.; ASVESTAS, P.; NIKITA, K. S.; MATSOPOULOS, G. K. Comparison of different global and local automatic registration schemes: An application to retinal images. In: **Lect. Notes Comput. Sc.** [S.l.: s.n.], 2004. p. 813–820. Citation on page 22.

KAUPPI, T. **Eye Fundus Image Analysis for Automatic Detection of Diabetic Retinopathy**. Phd Thesis (PhD Thesis) — Lappeenranta University of Technology, 2010. Citation on page 22.

KNOW, Y.; ADIX, M.; ZIMMERMAN, M.; PIETTE, S.; GREENLEE, E.; ALWARD, W.; ABRAMOFF, M. Variance owing to observer, repeat imaging, and fundus camera type on cup-to-disc ratio estimates by stereo planimetry. **J. Glaucoma**, v. 18, n. 4, p. 305–310, 2009. Citation on page 22.

KOHLER, T.; BUDAI, A.; KRAUS, M. F.; ODRSCILIK, J.; MICHELSON, G.; HORNEGGER, J. Automatic no-reference quality assessment for retinal fundus images using vessel segmentation. In: **Proc. CBMS IEEE**. [S.l.: s.n.], 2013. p. 95–100. Citation on page 62.

KOSOWSKY, J.; YUILLE, A. L. The invisible hand algorithm: Solving the assignment problem with statistical physics. **Neural networks**, Pergamon, v. 7, n. 3, p. 477–490, 1994. Citation on page 58.

LALIBERTÉ, F.; GAGNON, L.; SHENG, Y. Registration and fusion of retinal images: an evaluation study. **IEEE T. Med. Imaging**, v. 22, n. 5, p. 661–673, 2003. Citation on page 22.

LEE, J. A.; CHENG, J.; LEE, B. H.; ONG, E. P.; XU, G.; WONG, D. W. K.; LIU, J.; LAUDE, A.; LIM, T. H. A low-dimensional step pattern analysis algorithm with application to multimodal retinal image registration. In: **Proc. CVPR IEEE**. [S.l.: s.n.], 2015. p. 1046–1053. Citation on page 79.

LEE, S. **A study of ophthalmologic imaging distortion for retinal image registration and validation**. [S.l.]: ProQuest, 2008. Citation on page 41.

LEE, S.; REINHARDT, J. M.; CATTIN, P. C.; ABRAMOFFADE, M. D. Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model. **Medical Image Analysis**, v. 14, p. 539–549, 2010. Citation on page 23.

LEGG, P. A.; ROSIN, P. L.; MARSHALL, A. D.; MORGAN, J. E. Feature neighbourhood mutual information for multi-modal image registration: An application to eye fundus imaging. **Pattern Recogn.**, v. 48, n. 6, p. 1937–1946, 2015. Citation on page 34.

LEORDEANU, M. **Efficient Methods for Graph Matching and MAP Inference**. 2009. <https://sites.google.com/site/graphmatchingmethods/>. Citation on page 58.

LEORDEANU, M.; HEBERT, M. A spectral technique for correspondence problems using pairwise constraints. In: IEEE. **Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on**. [S.l.], 2005. v. 2, p. 1482–1489. Citation on page 58.

_____. A spectral technique for correspondence problems using pairwise constraints. In: **ICCV**. [S.l.: s.n.], 2005. v. 2. Citation on page 64.

LEORDEANU, M.; HEBERT, M.; SUKTHANKAR, R. An integer projected fixed point method for graph matching and map inference. In: **Advances in neural information processing systems**. [S.l.: s.n.], 2009. p. 1114–1122. Citation on page 58.

LEUTENEGGER, S.; CHLI, M.; SIEGWART, R. Y. Brisk: Binary robust invariant scalable keypoints. In: **IEEE I. Conf. Comp. Vis**. [S.l.: s.n.], 2011. p. 2548–2555. Citation on page 46.

LING, H.; OKADA, K. An efficient earth mover's distance algorithm for robust histogram comparison. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 29, n. 5, p. 840–853, 2007. Citations on pages 31 and 33.

LIU, Y.; WANG, X.-L.; WANG, H.-Y.; ZHA, H.; QIN, H. Learning robust similarity measures for 3d partial shape retrieval. **International Journal of Computer Vision**, Springer, v. 89, n. 2-3, p. 408–431, 2010. Citations on pages 33 and 34.

LOWE, D. G. Distinctive image features from scale-invariant keypoints. **Int. J. Comput. Vision**, v. 60, n. 2, p. 91–110, 2004. Citations on pages 36, 46, 59, 60, and 68.

M, D. C. B.; MOSES, C. J. Retinal image registration feature descriptors - a survey. In: **IEEE International Conference on Electronics and Communication Systems**. [S.l.: s.n.], 2014. p. 1–5. Citations on pages [23](#) and [65](#).

MAGNANTI, T. L.; AHUJA, R. K.; ORLIN, J. B. **Network Flows: Theory, Algorithms, and Applications**. [S.l.]: Pearson, 2014. Citation on page [41](#).

MATHWORKS, I. T. **MATLAB Image Processing Toolbox**. 2017. The MathWorks, Natick, MA, USA. Citation on page [60](#).

MATSOPOULOS, G. K.; MOURAVLIANSKY, N. A.; DELIBASIS, K. K.; NIKITA, K. S. Automatic retinal image registration scheme using global optimization techniques. **IEEE T. Inf. Technol. B.**, v. 3, n. 1, p. 47–60, 1999. Citation on page [36](#).

MCILROY, P.; ROSTEN, E.; TAYLOR, S.; DRUMMOND, T. Deterministic sample consensus with multiple match hypotheses. In: **Proc. BMVC**. [S.l.: s.n.], 2010. p. 1–11. Citations on pages [25](#) and [49](#).

MOSER, J. On the volume elements on a manifold. **Transactions of the American Mathematical Society**, JSTOR, p. 286–294, 1965. Citation on page [34](#).

MOTTA, D.; CASACA, W.; PAGLIOSA, P.; PAIVA, A. A user-friendly interactive framework for unsteady fluid flow segmentation and visualization. **Journal of Visualization**, Springer, p. 1–12, 2018. Citation on page [80](#).

MOTTA, D.; CASACA, W.; PAIVA, A. Fundus image transformation revisited: Towards determining more accurate registrations. In: **Proc IEEE Symp Comput Based Med Syst**. [S.l.: s.n.], 2018. Citations on pages [43](#), [49](#), [68](#), and [80](#).

MOTTA, D.; MATOS, L. de; SOUZA, A. C. de; MARCATO, R.; PAIVA, A.; CARVALHO, L. A. V. de. All-in-focus imaging technique used to improve 3d retinal fundus image reconstruction. In: **ACM. Proceedings of the 30th Annual ACM Symposium on Applied Computing**. [S.l.], 2015. p. 26–31. Citation on page [80](#).

MOTTA, D.; OLIVEIRA, M.; PAGLIOSA, P.; NONATO, L. G.; PAIVA, A. Exploratory segmentation of vector fields using multidimensional projection. In: **IEEE. Graphics, Patterns and Images (SIBGRAPI), 2015 28th SIBGRAPI Conference on**. [S.l.], 2015. p. 250–256. Citation on page [80](#).

MOTTA, D. A.; SERILLO, A.; MATOS, L. de; YASUOKA, F. M.; BAGNATO, V. S.; CARVALHO, L. A. 3d papillary image capturing by the stereo fundus camera system for clinical diagnosis on retina and optic nerve. In: **INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. Design and Quality for Biomedical Technologies VII**. [S.l.], 2014. v. 8936, p. 893614. Citation on page [81](#).

NIETO, E. G.; MOTTA, D.; NONATO, L. G. Semantically aware dynamic layouts. In: **IEEE. Graphics, Patterns and Images (SIBGRAPI), 2014 27th SIBGRAPI Conference on**. [S.l.], 2014. p. 220–226. Citation on page [81](#).

ONG, E. P.; LEE, J. A.; CHENG, J.; LEE, B. H.; XU, G.; LAUDE, A.; TEOH, S.; LIM, T. H.; WONG, D. W. K.; LIU, J. An augmented reality assistance platform for eye laser surgery. In: **Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. [S.l.: s.n.], 2015. p. 4326–4329. Citation on page [22](#).

PALRAJ, P.; VENILLA, I. A hybrid retinal image registration using mutual information. **Am. J. Appl. Sci.**, v. 10, n. 11, p. 1386–1391, 2013. Citation on page 35.

PALRAJ, P.; VENNILA, I. Retinal fundus image registration via blood vessel extraction using binary particle swarm optimization. **J. Med. Imaging Health**, v. 6, n. 2, p. 328–337, 2016. Citation on page 36.

PAPADAKIS, N. **Optimal Transport for Image Processing**. Phd Thesis (PhD Thesis) — Université de Bordeaux, 2015. Citation on page 30.

PELE, O.; WERMAN, M. Fast and robust earth mover's distances. In: IEEE. **Computer vision, 2009 IEEE 12th international conference on**. [S.l.], 2009. p. 460–467. Citations on pages 31 and 33.

PHILIP, S.; COWIE, L.; OLSON, J. The impact of the health technology board for scotland's grading model on referrals to ophthalmology services. **The British Journal of Ophthalmology**, v. 89, n. 1, p. 891–896, 2005. Citation on page 23.

PIECZYNSKI, J.; GRZYBOWSKI, A. Review of diabetic retinopathy screening methods and programmes adopted in different parts of the world. **European Ophthalmic Review**, v. 9, n. 1, p. 49–55, 2015. Citation on page 23.

POPLIN, R.; VARADARAJAN, A. V.; BLUMER, K.; LIU, Y.; MCCONNELL, M. V.; CORRADO, G. S.; PENG, L.; WEBSTER, D. R. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. **Nature Biomedical Engineering**, Nature Publishing Group, v. 2, n. 3, p. 158, 2018. Citation on page 21.

QUIGLEY, H. A.; BROMAN, A. T. The number of people with glaucoma worldwide in 2010 and 2020. **Brit. J. Ophthalmol.**, v. 90, n. 3, p. 262–267, 2006. Citation on page 21.

RABIN, J.; PEYRÉ, G.; COHEN, L. D. Geodesic shape retrieval via optimal mass transport. In: **Computer Vision–ECCV 2010**. [S.l.]: Springer, 2010. p. 771–784. Citations on pages 24, 33, and 34.

RAVUDU, M.; JAIN, V.; KUNDA, M. M. R. Review of image processing techniques for automatic detection of eye diseases. In: **Proc. ICST**. [S.l.: s.n.], 2012. p. 320–325. Citation on page 22.

REHMAN, T. ur; HABER, E.; PRYOR, G.; MELONAKOS, J.; TANNENBAUM, A. 3d nonrigid registration via optimal mass transport on the gpu. **Medical Image Analysis**, v. 13, n. 6, p. 931–940, 2009. Citations on pages 24, 32, and 33.

_____. 3D nonrigid registration via optimal mass transport on the GPU. **Med. Image Anal.**, v. 13, n. 6, p. 931–940, 2009. Citation on page 30.

RITTER, N.; OWENS, R.; COOPER, J.; EIKELBOOM, R. H.; SAARLOOS, P. P. van. Registration of stereo and temporal images of the retina. **IEEE Trans Med. Imaging**, v. 18, n. 5, 1999. Citation on page 36.

RUBNER, Y.; TOMASI, C.; GUIBAS, L. J. The earth mover's distance as a metric for image retrieval. **International Journal of Computer Vision**, Springer, v. 40, n. 2, p. 99–121, 2000. Citations on pages 24, 31, and 33.

RUECKERT, D.; ALJABAR, P. Nonrigid registration of medical images: Theory, methods, and applications. **IEEE Signal Processing Magazine**, v. 27, n. 4, p. 113–119, 2010. Citation on page 23.

SANTAMBROGIO, F. **Optimal Transport for Applied Mathematicians**. 1. ed. [S.l.]: Springer, 2015. ISBN 978-3-319-20827-5. Citations on pages 23 and 37.

SCHNABEL, J. A.; HEINRICH, M. P.; PAPIEŻ, B. W.; BRADY, J. M. Advances and challenges in deformable image registration: From image fusion to complex motion modelling. **Medical image analysis**, Elsevier, v. 33, p. 145–148, 2016. Citation on page 23.

SCHULTHEISS, T. E.; TOMÉ, W. A.; ORTON, C. G. It is not appropriate to “deform” dose along with deformable image registration in adaptive radiotherapy. **Medical physics**, Wiley Online Library, v. 39, n. 11, p. 6531–6533, 2012. Citation on page 22.

SERRADELL, E.; PINHEIRO, M. A.; SZNITMAN, R.; KYBIC, J.; MORENO-NOGUER, F.; FUA, P. Non-rigid graph registration using active testing search. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 37, n. 3, p. 625–638, 2015. Citations on pages 35, 59, 60, and 68.

SINKHORN, R. A relationship between arbitrary positive matrices and doubly stochastic matrices. **The annals of mathematical statistics**, JSTOR, p. 876–879, 1964. Citation on page 34.

SOLOMON, J. **Optimal Transport on Discrete Domains**. [S.l.], 2018. URL: <https://arxiv.org/abs/1801.07745>. Citations on pages 23 and 29.

SOLOMON, J.; GOES, F. de; PEYRÉ, G.; CUTURI, M.; BUTSCHER, A.; NGUYEN, A.; DU, T.; GUIBAS, L. Convolutional Wasserstein distances: efficient optimal transportation on geometric domains. **ACM T. Graphic.**, v. 34, n. 4, p. 1–11, 2015. ISSN 0730-0301. Citation on page 30.

SOLOMON, J.; GOES, F. de; STUDIOS, P. A.; PEYRÉ, G.; CUTURI, M.; BUTSCHER, A.; NGUYEN, A.; DU, T.; GUIBAS, L. Convolutional wasserstein distances: Efficient optimal transportation on geometric domains. **Preprint**, 2015. Citation on page 34.

SOLOMON, J.; RUSTAMOV, R.; GUIBAS, L.; BUTSCHER, A. Earth mover’s distances on discrete surfaces. **ACM Transactions on Graphics (TOG)**, ACM, v. 33, n. 4, p. 67, 2014. Citations on pages 33 and 34.

SOTIRAS, A.; DAVATZIKOS, C.; PARAGIOS, N. Deformable medical image registration: A survey. **IEEE transactions on medical imaging**, IEEE, v. 32, n. 7, p. 1153–1190, 2013. Citation on page 23.

STREHL, A.; GHOSH, J. Cluster ensembles—a knowledge reuse framework for combining multiple partitions. **J. Mach. Learn. Res.**, v. 3, n. Dec, p. 583–617, 2002. Citation on page 53.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. 1st. ed. [S.l.]: Springer-Verlag New York, Inc., 2010. Citation on page 52.

TSAI, C.-L.; LI, C.-Y.; YANG, G.; LIN, K.-S. The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence. **IEEE Transactions on Medical Imaging**, v. 29, n. 3, p. 636–649, 2010. Citations on pages 23, 59, and 65.

_____. The edge-driven dual-bootstrap iterative closest point algorithm for registration of multi-modal fluorescein angiogram sequence. **IEEE T. Med. Imaging**, v. 29, n. 3, p. 636–649, 2010. Citations on pages 36, 59, 60, and 68.

VARADHAN, R.; KARANGELIS, G.; KRISHNAN, K.; HUI, S. A framework for deformable image registration validation in radiotherapy clinical applications. **Journal of Applied Clinical Medical Physics**, v. 14, n. 1, p. 192–213, 2013. Citation on page 23.

VEDALDI, A.; FULKERSON, B. VLFeat: An open and portable library of computer vision algorithms. In: **Proc. ACM-MM**. [S.l.: s.n.], 2010. (MM '10), p. 1469–1472. Citation on page 60.

VIELMA, J. P. Mixed integer linear programming formulation techniques. **SIAM Rev.**, v. 57, p. 3–57, 2015. Citation on page 44.

VIERGEVER, M. A.; MAINTZ, J. B. A.; KLEIN, S.; MURPHY, K.; STARING, M.; PLUIM, J. P. A survey of medical image registration - under review. **Medical Image Analysis**, v. 33, p. 140–144, 2016. Citations on pages 22, 23, 51, and 53.

VILLANI, C. **Topics in optimal transportation**. [S.l.]: American Mathematical Soc., 2003. Citation on page 27.

_____. **Optimal transport: old and new**. [S.l.]: Springer, 2009. ISBN 978-3-540-71049-3. Citation on page 24.

_____. **Optimal transport: old and new**. [S.l.]: Springer, 2009. Citations on pages 24 and 34.

WANG, G.; WANG, Z.; CHEN, Y.; ZHAO, W. Robust point matching method for multimodal retinal image registration. **Biomed. Signal Proces.**, v. 19, p. 68–76, 2015. Citations on pages 36 and 79.

WANG, W.; SLEPČEV, D.; BASU, S.; OZOLEK, J. A.; ROHDE, G. K. A linear optimal transportation framework for quantifying and visualizing variations in sets of images. **International Journal of Computer Vision**, Springer, v. 101, n. 2, p. 254–269, 2013. Citations on pages 24, 32, and 33.

WANG, Z.; BOVIK, A. C. A universal image quality index. **IEEE Signal Proc. Letters**, v. 9, n. 3, p. 81–84, 2002. Citation on page 53.

WANG, Z.; BOVIK, A. C.; SHEIKH, H. R.; SIMONCELLI, E. P. Image quality assessment: from error visibility to structural similarity. **IEEE T. Image Proces.**, v. 13, n. 4, p. 600–612, 2004. Citation on page 54.

World Health Organization. **Visual impairment and blindness**. 2014. Available: <<http://www.who.int/mediacentre/factsheets/fs282/en/>>. Citation on page 21.

YANG, G.; STEWART, C.; SOFKA, M.; TSAI, C.-L. **Generalized Dual Bootstrap-ICP Algorithm**. 2007. <http://www.vision.cs.rpi.edu/gdbicp/>. Citation on page 60.

_____. Registration of challenging image pairs: Initialization, estimation, and decision. **IEEE T. Pattern Anal.**, v. 29, n. 11, p. 1973–1989, 2007. Citations on pages 36, 59, 60, and 68.

YANG, G.; STEWART, C. V.; SOFKA, M.; TSAI, C.-L. Registration of challenging image pairs: Initialization, estimation, and decision. **IEEE Trans. on PAMI**, v. 29, n. 11, p. 1973–1989, 2007. Citations on pages 23, 59, and 65.

YANG, Y.; HUANG, S. Retinal image mosaic base on genetic algorithm and automated blood vessel extracting approach. In: **7th World Congress on Intelligent Control and Automation**. [S.l.: s.n.], 2008. p. 7751–7756. Citation on page 36.

ZASLAVSKIY, M.; BACH, F.; VERT, J.-P. A path following algorithm for the graph matching problem. **IEEE T. Pattern Anal.**, v. 31, n. 12, p. 2227–2242, 2009. Citation on page 36.

ZASS, R.; SHASHUA, A. Probabilistic graph and hypergraph matching. IEEE, 2008. Citation on page 58.

ZHANG, Z.; SRIVASTAVA, R.; LIU, H.; CHEN, X.; DUAN, L.; KEE, D.; KWOH, C.; WONG, T.; LIU, J. A survey on computer aided diagnosis for ocular diseases. **BMC Med. Inform. Decis.**, v. 14, n. 1, p. 1–80, 2014. Citation on page 22.

ZHAO, X.; SU, Z.; GU, X. D.; KAUFMAN, A.; SUN, J.; GAO, J.; LUO, F. Area-preservation mapping using optimal mass transport. **Visualization and Computer Graphics, IEEE Transactions on**, IEEE, v. 19, n. 12, p. 2838–2847, 2013. Citations on pages 24, 33, and 34.

ZHOU, F.; De la Torre, F. Factorized graph matching. **IEEE T. Pattern Anal.**, v. 38, n. 9, p. 1774–1789, 2016. Citations on pages 58 and 64.

_____. **Factorized Graph Matching**. 2016. [Http://www.f-zhou.com/gm_code.html](http://www.f-zhou.com/gm_code.html). Citation on page 58.

ZHOU, W.; ZHANG, L.; XIE, Y.; LIANG, C. A novel technique for prealignment in multimodality medical image registration. **BioMed Res. Int.**, v. 2014, 2014. Citation on page 34.

ZHU, W.; LIANG, S.; WEI, Y.; SUN, J. Saliency optimization from robust background detection. In: **Proc. CVPR IEEE**. [S.l.: s.n.], 2014. p. 2814–2821. Citation on page 47.

ZITOVÁ, B.; FLUSSER, J. Image registration methods: A survey. **Image and Vision Computing**, v. 21, n. 11, p. 977–1000, 2003. Citations on pages 23, 51, and 53.