Multimodal Macula Mapping by Deformable Image Registration

P. Baptista
Center of New Technologies for Medicine (CNTM) from the Association for Innovation and Biomedical Research on Light and Image (AIBILI), Coimbra, Portugal
J. Ferreira
Institute of Systems and Robotics (ISR), Coimbra, Portugal
R. Bernardes
Center of New Technologies for Medicine (CNTM) from the Association for Innovation and Biomedical Research on Light and Image (AIBILI), Coimbra, Portugal
J. Dias
Institute of Systems and Robotics (ISR), Coimbra, Portugal
Faculty of Sciences and Technology, University of Coimbra, Coimbra, Portugal
J. Cunha-Vaz
Center of New Technologies for Medicine (CNTM) from the Association for Innovation and Biomedical Research on Light and Image (AIBILI), Coimbra, Portugal
Institute of Biomedical Research on Light and Image (IBILI), Faculty of Medicine, University of Coimbra, Coimbra, Portugal

ABSTRACT: This work aims to present the applied methods for the registration of different imaging modalities of the human macula. Multimodal image registration is herewith considered the process of overlaying two or more images of the same eye, taken from different viewpoints, at different times, with different sizes and fields–of–view and using different sensors, thus imaging different retinal structures. Additionally, the registration of retinal images from *in vivo* patients using scanning systems has to deal with intrinsic image distortions due to saccadic eye movements. In order to correct for differences in eye fundus scanned angles, image sizes and feature expressions in different modalities, a *feature vector* was built for each modality embodying complementary information from the image alone.
KEYWORDS: Multimodal, Registration, Deformable Registration

1 INTRODUCTION

Multimodal macula mapping is the combination of a variety of diagnostic imaging modalities to examine the macular region in order to obtain information on its structure and function. The macula is located in the posterior pole of the retina and is responsible for detailed and color vision, thus constituting an important area for human vision, as any macular alteration will, sooner or later, affect visual acuity (Bernardes, Lobo, and Cunha-Vaz 2002). Currently, a wealth of retina imaging modalities exists, either on daily clinical practice or research environments, such as color fundus photography, red–free fundus photography, fluorescein angiography, indocyanine green angiography, optical coherence tomography, retina flowmeter, fundus autofluorescence, leakage analysis, multifocal electroretinography, etc. The potential for multimodal macular mapping was demonstrated in (Bernardes, Lobo, and Cunha-Vaz 2002).

Laliberté and Gagnon stated in (Laliberté and Gagnon 2003) that ophthalmology is one of the medical areas where diagnosis implies the analysis of a large number of images. This is true even for a single eye followed during a short period of time due to the number of complementary examination procedures in use for pathologies like diabetic retinopathy (DR) or age–related macular degeneration (AMD), to name but a few. Duncan and Ayache, in (Duncan and Ayache 2000), gave an overview of medical image analysis discussing, among other topics, methodologies for image matching/registration. On their comments on multimodal image registration we can find the problems due to images being formed by various

imaging modalities producing different types of information and often at different spatial resolutions.

Published work on multimodal retinal image registration deals mostly with images acquired using area sensors, with which the whole image is acquired at once (Laliberté and Gagnon 2003; Zana and Klein 1999; Matsopoulos, Asvestas, Mouravliansky, and Delibasis 2004), while available retinal imaging modalities encompass a number of different sensors and sensor types.

As the use of confocal scanning laser ophthalmoscopes (CSLO) (Webb, Hughes, and Delori 1987) becomes more commonplace, a new problem arises for the registration of this sort of images, more precisely the intrinsic image distortions due to saccadic eye movements, which in turn can be of two kinds: voluntary (e.g., when the eyes move from one fixation point to another as in reading) and involuntary (e.g., due to following the flying spot that illuminates the eye fundus). These effects in the image will obviously depend on the amount of time for the full scan, e.g., 400 ms for the Zeiss CSLO used in (Bernardes, Lobo, and Cunha-Vaz 2002) to 1.6 s for the CSLO used in this work. This suggests the need for deformable image registration procedures to correct for local deformations as opposed to the statement of global deformations on ophthalmic images by Laliberté and Gagnon in (Laliberté and Gagnon 2003) and the statement of Zana and Klein in (Zana and Klein 1999), where the authors restricted their work to central images containing the macula, papila and temporal vessels to limit deformations. In (Pinz, Bernögger, Datlinger, and Kruger 1998), Pinz *et al.* published a multimodal imaging system based on a scanning laser ophthalmoscope without taking into consideration the distortions herein considered. Nevertheless, the imaging device used in their work was not specified besides image size and scanned angle, 760 × 430 pixels and 40°, respectively. The system may, therefore, be fast enough to produce results similar to an area sensor and thus not showing local distortions.

In our work we will focus on two different types of light detectors – area sensors, either of color or monochromatic type, and single–point sensors (e.g., photodiodes) – to acquire *in vivo* $2D$ morphology and/or functional information of the human retina. Moreover, we will address the problem of registering different modalities based on different sensors and covering different fields–of–view (FOV), e.g., 20° high–speed fluorescein angiography centered on the fovea (performed by a CSLO) to a 50° color fundus photograph. Another major difference to published literature is related to the fact that being centered on the foveal area and using small FOV, we are faced with a low number of retinal structures. Furthermore, the modalities herewith considered gather a different level of details for this particular region of the human eye, therefore not being possible to apply the methods presented in (Laliberté and Gagnon 2003; Zana and Klein 1999; Matsopoulos, Asvestas, Mouravliansky, and Delibasis 2004).

In the remaining of this text we will present a set of methodologies that allow automatic registration of images from different imaging modalities (e.g., retinographies and fluorescein angiographies) and using different sensor types (area and single–point sensors).

2 METHODS

2.1 *Modalities*

One of the most widely used retinal imaging modality is the retinography (color fundus photograph), either on film or digital format. Currently, a large number of digital color fundus photography devices exist, producing images ranging in size from thousands– to mega–pixels. In this work, a Zeiss Fundus Camera system, model FF450, was used to acquire 768 × 576 pixels RGB images of 50° FOV, centered on the fovea, thus imaging the whole macula, optic disc and temporal retinal arcades.

The other modality herewith considered is the fluorescein angiography, i.e., a fundus photograph taken after the administration of a dye (sodium fluorescein, NaFl, $C_{20}H_{10}O_5Na_2$). Fluorescein angiographies may be performed using either a regular camera, similar to the ones used for retinographies, or using a CSLO, thus being named high–speed fluorescein angiographies or SLO.FA. Using a series of fluorescein angiographies taken with a CSLO in different confocal planes, it is possible to compute the leakage of fluorescein into the vitreous and/or the permeability of the blood–retinal barrier to fluorescein (Lobo, Bernardes, Santos, and Cunha-Vaz 1999; Bernardes, Dias, and Cunha-Vaz 2005). In this work, a Heidelberg Retina Angiograph was used to acquire 20° SLO.FA of 256 × 256 pixels (maximum).

2.2 *Features*

The features available for each modality herein considered are the fovea, the optic disc and the vascular network for retinographies, and the fovea and a small set of the vascular network for the SLO.FA.

The retinography is pre–processed to be converted into a gray–scale image, where the fovea appears as a dark region, the optic disc appears as a bright area and the vessels, veins and arteries, are darker than the surrounding tissue, i.e., the background.

Vascular network expression in fluorescein angiographies is dependent on the time after injection, i.e., the amount of time elapsed since the dye was injected into the patient blood-stream. After injection, arteries become brighter as fluorescein starts appear-

ing in the retinal circulation and veins are still dark. Thereafter, also veins end up getting brighter as fluorescein is collected after microcirculation. The foveal area is normally kept darker than background. For the time frame of interest for our application, 10 to 20 minutes after dye injection, the fovea is characterized by a dark region and the vascular network by bright vessels.

As different modalities may represent the same feature in different ways and cover different areas of the eye fundus using different image sizes, a feature vector of the type $f_{modality} = \langle$*modality, color model, field–of–view, image size, optic disc diameter, time–after–injection,* $\ldots\rangle$ is created for each modality, encompassing additional information from the image alone.

As modalities herein considered cover different areas of the eye fundus, 20° and 50°, a coarse registration approach is mandatory by computing the location of common features for both modalities. Therefore, for each modality, the location of the fovea is determined by computing their cross–correlation with an inverted Gaussian (1), with σ experimentally determined ($k = 1$).

$$g(i,j) = k\left[1 - \frac{1}{2}exp\left(\frac{-(i^2+j^2)}{2\sigma^2}\right)\right] \quad . \qquad (1)$$

For each modality, the vascular network is detected through skeletonization after vascular tree segmentation. This segmentation was achieved resorting to contourlets (Do 2003), an efficient directional multiresolution image representation which provides the usual advantages of wavelet analysis, adding anisotropy and directionality properties which the latter lacks, presenting the basis to a more robust method for smooth contour segmentation. In other words, it provides the same level of performance of wavelets using fewer coefficients for this type of segmentation (Do 2003) — this fact was confirmed in our experimental studies.

On the other hand, comparatively to other techniques based on radically different approaches for segmentation, such as morphology or differential geometry, the contourlet method has proved in our experimental studies to be more robust in the presence of noise in images, or for low resolution/small FOV angle modalities (e.g. 20° SLO.FA images) where even apparently high signal–to–noise ratios greatly influence the vessel detection procedure.

2.3 *Global Rigid Registration*

Since modalities have different image sizes and FOV and the registration approach will not be based on computing a transformation matrix (affine or perspective) based on a set of corresponding landmarks, it is mandatory to have all modalities at the same scale, with each pixel representing the same FOV. In order to avoid the creation of artifacts, instead of interpolation we followed the approach of decimating the higher resolution modality to the resolution of the one having the lower number of pixels per degree of the eye fundus. Therefore, as the retinography presents a resolution of 14.5 $pixels/degree$ and the SLO.FA presents 12.8 $pixels/degree$, the retinography was decimated to the latter resolution at the pre–processing stage.

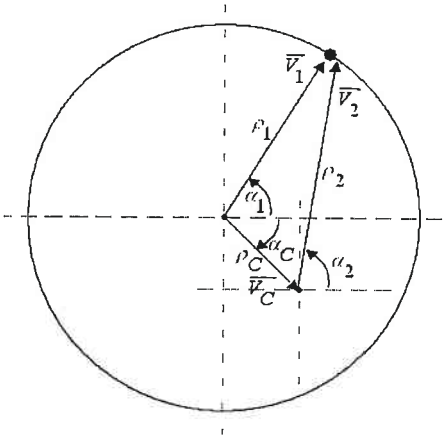

Figure 1: Vectorial representation of the translation of the centre of the fovea.

Having extracted both the location of the fovea and the retinal vascular tree for each modality, it is possible to compute the rotation between modalities by computing the translation in polar coordinates of the extracted retinal vascular network.

As the location of the fovea in one modality may be estimated with some displacement relatively to the other modality, this displacement influences the rigid registration both on the translational part as well as in the rotational part, as seen in Figure 1. This relative displacement produces a distortion in the representation of the vascular network in polar coordinates, both in the angle and distance to the reference, i.e., $\alpha_1 \rightarrow \alpha_2$ and $\rho_1 \rightarrow \rho_2$ (Fig. 1), being these effects dependent on the relative distances between references (relative positions of detected foveas) and the point to be mapped into polar coordinates.

By computing block translations in polar coordinates, it is possible to compute both α_C and ρ_C (Fig. 1), since for α_C and $\alpha_C + \pi/2$ there is a zero shift in the angle and the maximal shift in ρ_C, being it equal to half the difference between ρ_{α_C} and $\rho_{\alpha_C+\pi/2}$. Both ρ_C and α_C can be determined by a minimization algorithm as $\rho_2 = f(\rho_1, \rho_C, \alpha_1, \alpha_C)$ (2).

$$\rho_2 = \sqrt{\rho_1^2 + \rho_C^2 - 2\rho_1\rho_C\cos(\alpha_1 - \alpha_C)} \quad . \qquad (2)$$

This procedure allows to simultaneously correct

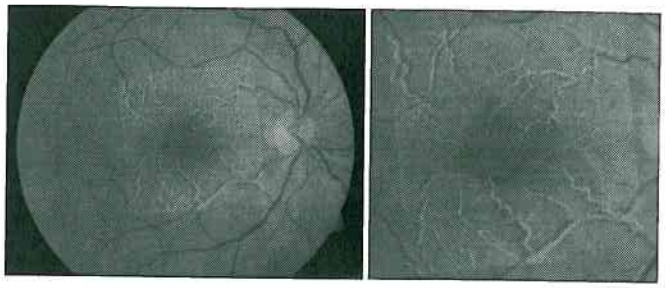

Figure 2: Result of rigid registration showing room for improvement.

differences in the computed fovea between images (modalities) as well as correct the respective distortions in polar coordinates, thus making it possible to compute the rotation between modalities. It is now possible to perform global rigid registration between modalities which represents the first step for deformable registration, as seen in Figure 2, i.e. to compute a 3×3 transformation matrix $\mathbf{P}$,

$$P = T(x_{retino}, y_{retino}) \times R(\theta) \times T(-x_{slo.fa}, -y_{slo.fa}) \quad (3)$$

where (x_{retino}, y_{retino}) and $(x_{slo.fa}, y_{slo.fa})$ (translation parameters) are the homogenous fovea coordinates on both modalities (after correction) and θ is the estimated rotation angle between modalities.

2.4 *Deformable Registration*

After the initial approach using global rigid registration, it is possible to fine–tune the registration by considering the SLO.FA as a set of overlapping windows ($W_{SLO.FA}$) of size $M \times M$ centered on pixel (x_i, y_i), corresponding to pixel (x_f, y_f) on the retinography, i.e. $\mathbf{X_f} = \mathbf{P}\mathbf{X_i}$, with $\mathbf{X_f} = [x_f, y_f, 1]^T$, $\mathbf{X_i} = [x_i, y_i, 1]^T$ and $\mathbf{P}$ as computed in (3).

A W_{RETINO} window of size $N \times N$ centered on (x_f, y_f) (on the retinography) constitutes the search space where a match will be sought for the $W_{SLO.FA}$ window. This search is made by computing a similarity measurement using the partitioned intensity uniformity (PIU) defined by (4). For details see (Hill and Batchelor 2001).

$$PIU_A = \sum_b \frac{n_b}{N} \frac{\sigma_A(b)}{\mu_A(b)} \quad . \quad (4)$$

For rotation angles below a given threshold (θ_{thres}), the matching is performed by ignoring the rotation, while for angles above θ_{thres} the SLO.FA is previously rigidly registered.

For each $(W_{RETINO}, W_{SLO.FA})$ window pair a translation is therefore computed on top of the previous transformation $\mathbf{P}$, thus being computed a transformation $\mathbf{P'}$ for each $W_{SLO.FA}(i,j)$ window position

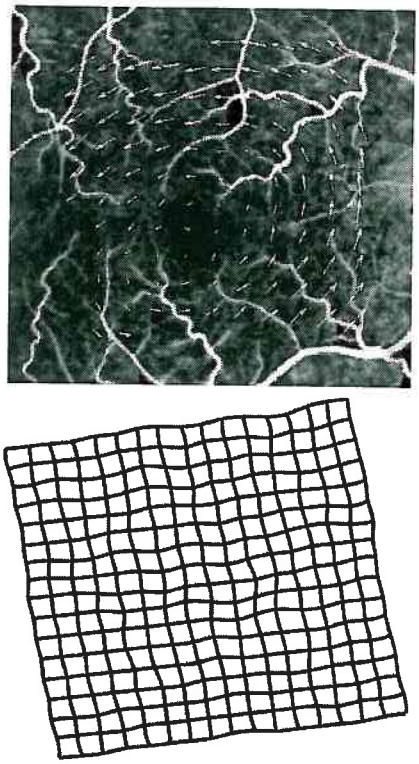

Figure 3: The deformation field calculated using equations above (scaled for better clarity) over the SLO.FA (top) and the grid deformation (bottom).

as given by

$$\mathbf{P'}_{i,j} = \begin{bmatrix} 1 & 0 & \Delta x(i,j) \\ 0 & 1 & \Delta y(i,j) \\ 0 & 0 & 1 \end{bmatrix} \mathbf{P} \quad , \quad (5)$$

where i, j represents the $W_{SLO.FA}$ window index.

A set of control points, computed by (5), allows to establish two thin-plate splines (TPS), one for the displacement in the x-direction and the other for the y-direction. Figure 3 shows the computed deformation field.

3 RESULTS AND CONCLUSIONS

The final result achieved with the developed methodology as explained in this text is shown on Figure 4. The quality of the registration was qualitatively assessed as no quantitative algorithm was implemented yet. Nevertheless, it is now clear the need for the deformable registration for multimodal macula mapping when considering SLO–like modalities, and that the approach followed produces registration with the necessary accuracy for the current purposes.

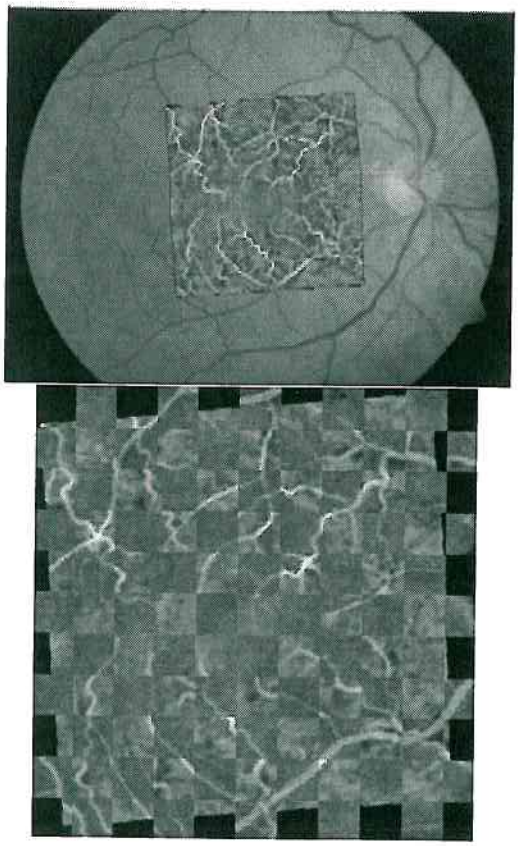

Figure 4: Top: The SLO.FA (gray-scale portion) shown over the full retinography (color). Bottom: Gray–scale chessboard–like representation of both modalities (only the common portion).

REFERENCES

Bernardes, R., J. Dias, and J. Cunha-Vaz (2005, January). Mapping the human blood-retinal barrier function. *IEEE Trans. Biomed. Eng. 52*(1), 106–116.

Bernardes, R., C. Lobo, and J. Cunha-Vaz (2002, November–December). Multimodal macula mapping: A new approach to study diseases of the macula. *Surv Ophthalmol 47*(6), 580–589.

Do, M. (2003). Contourlets and sparse image expansions. In *Wavelets: Applications in Signal and Image Processing X. Proceedings of the SPIE*, pp. 560–570.

Duncan, J. and N. Ayache (2000, January). Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis an Machine Intelligence 22*(1), 85–106.

Hill, D. and P. Batchelor (2001). Registration methodology: Concepts and algorithms. In J. Hajnal, D. Hill, and D. Hawkes (Eds.), *Medical Image Registration.*, The Biomedical Engineering Series, pp. 39–70. CRC Press.

Laliberté, F. and L. Gagnon (2003, May). Registration and fusion of retinal images—an evaluation study. *IEEE T Med Imaging 22*(5), 661–673.

Lobo, C., R. Bernardes, F. Santos, and J. Cunha-Vaz (1999, May). Mapping retinal fluorescein leakage with confocal scanning laser fluorometry of the human vitreous. *Arch Ophthalmol 117*(5), 631–637.

Matsopoulos, G., P. Asvestas, N. Mouravliansky, and K. Delibasis (2004, December). Multimodal registration of retinal images using self organizing maps. *IEEE T Med Imaging 23*(12), 1557–1563.

Pinz, A., S. Bernögger, P. Datlinger, and A. Kruger (1998, August). Mapping the human retina. *IEEE T Med Imaging 17*(4), 606–619.

Webb, R., G. Hughes, and F. Delori (1987, April). Confocal scanning laser ophthalmoscope. *Appl. Opt. 26*(8), 1492–1499.

Zana, F. and J. C. Klein (1999, May). A multimodal registration algorithm of eye fundus images using vessels detection and hough transform. *IEEE T Med Imaging 18*(5), 419–428.

CompIMAGE

20-21 OCTOBER 2006

COMPUTATIONAL MODELLING OF OBJECTS REPRESENTED IN IMAGES FUNDAMENTALS, METHODS AND APPLICATIONS

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

SUPPORTING ORGANIZATIONS

Centro Internacional de Matemática
Associação Portuguesa de Mecânica Teórica, Aplicada e Computacional
Faculdade de Engenharia da Universidade do Porto