DISTORTION FREE REGISTRATION BETWEEN MULTIFOCAL ERG AND RETINAL LEAKAGE ANALYZER

Rui Bernardes*, João Filipe Ferreira*,**, Pedro Baptista*, Ana Sebastião*,***, Jorge Dias** and José Cunha-Vaz*,***

* Center of New Technologies for Medicine, Association for Innovation and Biomedical Research on Light and Image, Coimbra, Portugal
** Institute of Systems and Robotics, University of Coimbra, Coimbra, Portugal
*** Institute of Biomedical Research on Light and Image, Faculty of Medicine of the University of Coimbra, Coimbra, Portugal

rcb@aibili.pt

Abstract: The aim of the present work is the registration of two functional imaging modalities while simultaneously correcting for the intrinsic distortions present on both. Two recently introduced imaging modalities are used in this work, one that measures the retina response to blue light stimuli and another that measures the blood-retinal barrier function. A deformable registration procedure was developed, comprising a model for saccadic eye movements which allows for an important reduction in the number of parameters which need to be estimated. Finally, simultaneous correction for the intrinsic distortions was achieved through resorting to an additional imaging modality, thus resulting in practice in the registration of three different modalities, the former two of the functional type and the latter of the morphological type.

Introduction

Multimodal macula mapping, the combination of a variety of diagnostic imaging modalities to examine the macular region, aims to obtain information on macular 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 [1]. 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, leakage analysis, multifocal electroretinography, to name but a few. The potential for multimodal macular mapping was demonstrated in [1].

As the use of confocal scanning laser ophthalmoscopes (CSLO) [2] 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 assembled image will obviously depend on the amount of time for the full scan, e.g., 400 ms for the Zeiss CSLO used in [1] to 1.6 s at the fastest scanning rate for the CSLO used in this work as the source for the leakage analysis modality [3]. 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 [4] and the statement of Zana and Klein in [5], where the authors restricted their work to central images containing the macula, papila and temporal vessels to limit deformations. In [6], 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, 768×576 pixels and 40°, respectively. The system may, therefore, be fast enough to produce results similar to an area sensor and thus not showing visible local distortions.

The aim of the present work is the registration of two functional imaging modalities using a color fundus photograph (retinography) to correct for intrinsic distortions — the multifocal electroretinography (mfERG), which measures the retina response to light stimuli, and the retinal leakage analyzer (RLA), which measures the leakage of fluorescein from the blood-stream into the human vitreous. Both of these modalities are performed *in vivo* and allow the detection of different types of early changes in the human retina in diseases such as diabetes.

Materials and Methods

Multifocal ERG is widely used in the detection of local functional deficits of the retina. It is most useful in the detection of the topographical distribution of diseases in many inherited retinal degenerations. The method introduced by Sutter and Tran in [7] is based on the *m-sequence* stimulation technique and allows for the simultaneous measurement of the ERG activity of many retinal locations [8].

However, the position of the stimulus on the retina

ISSN: 1727-1983

and the stability of fixation are usually not directly accessible. Although fixation is not considered a major issue in most clinical situations, viewing and recording the actual fundus position during stimulation is important for many non-standard applications, like the analysis of eccentric regions or the evaluation of transplants and retinal prostheses. Recently, a method was developed that uses a scanning laser ophthalmoscope for a combined stimulation and imaging of the retina [8].

The Heidelberg Retina Angiograph II (HRA, Heidelberg Engineering, Germany) is a CSLO that provides a blue laser (488 nm) for stimulus pattern generation and an infrared laser (835 nm) for fundus visualization. This system was used connected to a RETIScan device (RETIScan, Roland Consulting, Germany) to allow the production of stimuli in a 15° field-of-view (FOV) centered on the fovea and the registration of the eye fundus reference image of 384×384 pixels.

On the other hand, the RLA is a technique that computes fluorescein leakage from the blood-stream into the human vitreous, *in vivo*, by means of three-dimensional volumetric information with simultaneous imaging of the retina, thus establishing a direct correlation between retinal leakage and retinal morphology [3]. This leakage information is in fact information about the blood-retinal barrier permeability to sodium fluorescein.

As shown in [3], two types of information are extracted from the 3D fluorescein distribution: the fundus reference, a fluorescein angiography similar image, and; the leakage map, i.e., the functional information on the blood-retinal barrier function. These two sets of data share a common reference, have a pixel-to-pixel relation and cover the 20° FOV producing a leakage map/fundus reference of 256×256 pixels (maximum).

As mentioned above, a retinography is used in this work so as to correct for the intrinsic deformations present on both the mfERG and RLA. Its use is due to the widespread availability of this modality and to the fact that, contrary to the other two modalities in question, it acquires the entire image simultaneously, therefore not yielding distortions resulting from saccadic eye movements. This modality consists of an RGB color photograph, 50° FOV and 768×576 pixels.

Three assumptions are made in this work:

First assumption All the modalities herewith considered contain the foveal area.

Second assumption Pixel resolution (i.e. the number of pixels per degree on the eye fundus) is known for each modality.

Third assumption The deformation consists of local translations (saccadics) on top of global rigid transformation.

The first assumption comes directly from this work's application area. The second assumption, besides small differences due to misfocusing, holds true both for the mfERG and RLA, while for the retinography there are a set of fixed setups which are easily configured as inputs to our processes. Finally, the third assumption is based on acquired experience.

Since the retinography imaging modality was chosen as the fundus reference, the registration of the two functional imaging modalities to one another is achieved by using the coordinate system established by the retinography. Therefore, two registrations are in fact performed: the mfERG to the retinography, and the RLA to the retinography. In other words, the mfERG and the RLA are registered through their projection onto the coordinate system established by the retinography.

In order to have a single channel fundus reference for the retinography, a principal component analysis (PCA) using its RGB channels was performed and the first principal component (associated to the highest variance) used as the fundus reference, since it captures the largest amount of information spread over the three color channels.

Each registration now consists of computing a mapping between each of the mfERG and RLA fundus images to the reference established above. In formulating this problem, we can consider the image domain to be the area of $\mathbb{R}^2$ occupied by the retinography onto which both the mfERG and RLA are to be mapped. The reference coordinate system defined by the retinography and the coordinate systems defined by each of the other two modalities are related by a generic parametric mapping defined by

$$x^r = f(\Gamma^i, x^i) \tag{1}$$

where f is the mapping function and Γ^i are the parameters for the mapping between x^i and x^r, respectively the pixel coordinates in each of the modalities (i) and in the reference coordinate system (r). We need therefore to determine all the unknown parameters $(\gamma_1^i, \gamma_2^i, \ldots, \gamma_n^i)$ composing Γ^i, which brings the modalities involved into registration.

Given our deformation model (third assumption), it is possible to define $\mathbf{P}_{j,k}^i$ for each pixel coordinate (j,k) so that

$$\mathbf{P}_{j,k}^i = \mathbf{S}_{j,k}\mathbf{P}_{global}^i \tag{2}$$

where $\mathbf{P}_{j,k}^i$ is the *local* transformation matrix, $\mathbf{P}_{global}^i$ accounts for the *global* transformation and $\mathbf{S}_{j,k}$ accounts for saccadics. Therefore equation (1) becomes (3)

$$x^r = \mathbf{P}_{j,k}^i x^i \tag{3}$$

Taking into consideration the differences in the modalities involved, the partitioned intensity uniformity (PIU), as described in [9], was selected as the similarity measure to be used.

More precisely, using (1) by considering two generic images A and B for which a transformation $A = f(\Gamma_{j,k}^B, B)$ is defined so as to obtain domains Ω_a^f and Ω_b^f, with a and b being intensity values or ranges respectively taken from each of these images, and x^B being pixel coordinates in

ISSN: 1727-1983

image B and $x^A = f(\Gamma^B_{x^B}, x^B)$ being the transformed pixel coordinates from image B to image A, we have

$$PIU(A,B) = \sum_b \frac{n_b}{N} \frac{\sigma_A(b)}{\mu_A(b)} \quad . \tag{4}$$

where

$$n_b = \sum_{\Omega_b^f} 1 \tag{5}$$

$$\mu_{\mathbf{A}}(b) = \frac{1}{n_b} \sum_{\Omega_b^f} \mathbf{A}\left(x^A\right) \tag{6}$$

$$\sigma_{\mathbf{A}}(b) = \frac{1}{n_b} \sum_{\Omega_b^f} \left(\mathbf{A}\left(x^A\right) - \mu_{\mathbf{A}}(b)\right)^2 \tag{7}$$

These equations imply that the similarity between reference image B and image A mapped through f increases as the PIU value decreases.

If we denote the similarity between modality i mapped through f and the retinography as PIU^i, our aim is therefore to determine $\Gamma^i_{j,k}$ by minimizing

$$\hat{\Gamma}^i_{j,k} = \arg\min_{\Gamma^i_{j,k}} PIU^i(I^r(f(\Gamma^i_{j,k}, x^i)), I^i(x^i)) \tag{8}$$

which, given the equations above, is equivalent to minimizing

$$\hat{\Gamma}^i_{j,k} = \arg\min_{\Gamma^i_{j,k}} PIU^i(I^r(\mathbf{P}^i_{j,k} x^i), I^i(x^i)) \tag{9}$$

where I^r is a pixel on the retinography with coordinates given by (3). Therefore, for each pixel on modality i one needs to determine a window of size $N \times N$ centered on (j,k), to project this window onto the reference coordinate system, compute the PIU^i similarity value between the original and projected windows and determine the parameters $\Gamma^i_{j,k}$ that yield the best matching. In other words, one needs to compute a transformation matrix of the type

$$\mathbf{P}^i = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 \\ \gamma_4 & \gamma_5 & \gamma_6 \\ \gamma_7 & \gamma_8 & 1 \end{bmatrix} \quad , \tag{10}$$

for each pixel (j,k). This transformation is therefore of the rigid, affine or perspective type depending on the values for γ_1 to γ_8.

Since P^i in (10) is defined as in (2), the previous equation reduces to

$$\mathbf{P}^i = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 \\ \gamma_4 & \gamma_5 & \gamma_6 \\ 0 & 0 & 1 \end{bmatrix} \quad , \tag{11}$$

Indexes for pixel coordinates in (10) and (11) were omitted for clarity.

To deal with scalings among different modalities, images were prescaled to the lowest resolution available, which corresponds to the RLA. One should remember that our initial goal was to register the mfERG and RLA modalities and correct for their intrinsic distortions. Therefore, reducing the retinography and mfERG image resolutions to the RLA resolution should not be considered a drawback. Our initial problem in now reduced to finding a rigid registration (translation plus rotation), followed by local translations, for each modality i. Therefore, firstly we need to compute parameters t^i_x, t^i_y and θ^i that define P^i_{global}, i.e., the global rigid registration (2).

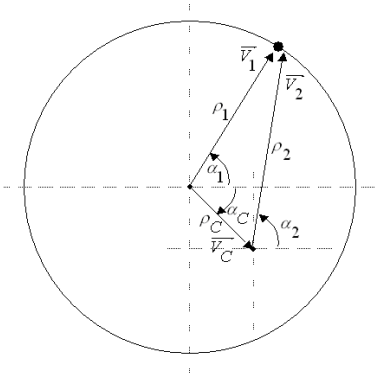

Figure 1: Vectorial representation of the translation of the center of the fovea.

Initial estimates for t^i_x and t^i_y, respectively t^i_{x0} and t^i_{y0}, can be computed directly from the images without the need for any optimization procedure. More specifically, since the fovea is present in each of the modalities, its location at each modality can be estimated by a cross-correlation process between the respective fundus reference and an inverted Gaussian as in [10]. Therefore, t^i_{x0} and t^i_{y0} are computed as the difference between the estimated fovea location at the retinography (Φ^r_x, Φ^r_y) and the respective modality i, i.e.,

$$\begin{cases} t^i_{x0} = \Phi^r_x - \Phi^i_x \\ t^i_{y0} = \Phi^r_y - \Phi^i_y \end{cases} \tag{12}$$

Some small differences in the detected fovea coordinates at each modality might occur in this phase — this will be corrected later on.

In order to compute the rotation θ^i, the vascular network for each modality was detected through skeletonization after vascular tree segmentation. This segmentation was achieved resorting to contourlets [11], 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 [11]. 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 shown itself during the course of our work to be more robust in the presence of noise or for low resolution/small FOV angle modalities, where even apparently

ISSN: 1727-1983

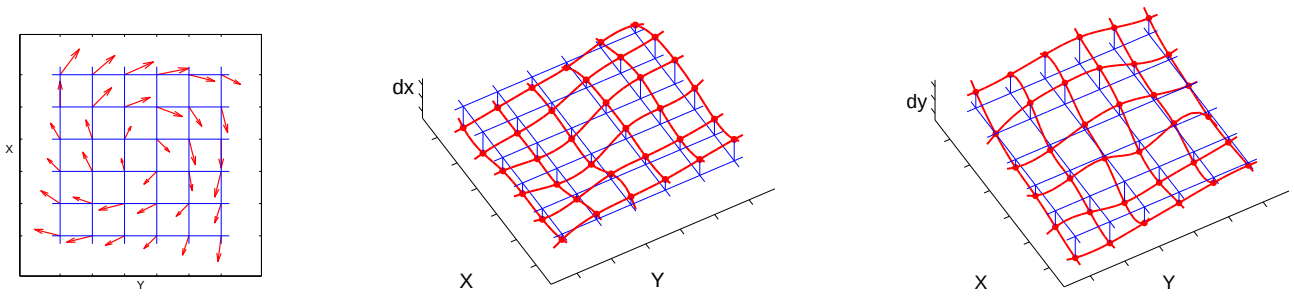

Figure 2: Deformation vector-field example (left), with resulting displacement in x (middle) and in y (right).

high signal-to-noise ratios greatly influence the vessel detection procedure.

Vessel centerlines were then plotted in polar coordinates relatively to the computed fovea location in each modality so to compute the rotation between modalities, i.e., translation in polar coordinate along the θ axis.

Since both the mfERG and RLA modalities have small fundus reference images, all the image space should be considered for representation in polar coordinates. Therefore, the aforementioned differences in the detected position of the fovea will have as a side-effect the distortion of vessel centerlines in the polar representation. As can be seen in figure 1 on the preceding page the same physical location can be represented in polar coordinates by (ρ_1, α_1) or (ρ_2, α_2), depending on the viewpoint, and the distortion increases with the decrease of the radius.

Vector $\vec{v_c}$, as shown in figure 1, establishes the relative locations of the detected fovea in the different modalities. As ρ_2 is given by (13), it is possible to compute $\vec{v_c} = (\rho_c, \alpha_c)$ by fitting, through an optimization procedure, ρ_2 and α_2 as in (13) to the computed displacements, by a phase cross-correlation process, in ρ and α (polar coordinates, radius and angle, respectively) of windows of size $M \times M$. For details, please refer to [12].

$$\rho_2 = \sqrt{\rho_1^2 + \rho_C^2 - 2\rho_1\rho_C \cos(\alpha_1 - \alpha_C)} \ . \qquad (13)$$

Once $\vec{v_c}$ is computed, its projections onto the $x-$ and $y-$axis are used to update the initial estimate for the translation between modalities, i.e.,

$$\begin{cases} t_x^i = t_{x0}^i + < \vec{v_c}, \hat{x} > \\ t_y^i = t_{y0}^i + < \vec{v_c}, \hat{y} > \end{cases} \qquad (14)$$

where $\hat{x}$ and $\hat{y}$ are respectively the unit vectors on the $x-$ and $y-$axis, and $< \cdot >$ represents the inner product operation.

It is now possible to re-compute the new polar representation of vessel centerlines of modality i and to determine the rotation θ^i by a phase cross-correlation process, without the influence of deformations. Rigid registration between modality i and the retinography is now given by $\mathbf{P}^i$ as (15),

$$\mathbf{P}^i = \mathbf{H}^f \, \mathbf{R}_{\theta \mathbf{i}} \, \mathbf{H}^i \qquad (15)$$

where $\mathbf{H}^i$ translates the *corrected* fovea coordinates of modality i to the origin $(\mathbf{0})$, $\mathbf{R}_{\theta \mathbf{i}}$ rotates modality i of θ^i and $\mathbf{H}^f$ translates $\mathbf{0}$ to the coordinates of the fovea on the reference coordinate system — (Φ_x^r, Φ_y^r) — as in (12).

While this is the established theoretical transformation, only for θ^i over a threshold θ_{thres} is the rotation applied in practice, otherwise $\mathbf{R}_{\theta \mathbf{i}}$ becomes the identity matrix and $\mathbf{P}^i$ reduces to $\mathbf{P}^{\mathbf{i}} = \mathbf{H}^f \mathbf{H}^i$.

At this stage and following the established deformation model, non-linear distortions can be correct by block matching, that is, searching for the best fit in the PIU metric sense (i.e., lower PIU values correspond to best fits) of overlapping windows $W_{m,n}^i$ (where (m,n) are the coordinates of each window center) in modality i to windows W^r (retinography) centered on x^r given by (15).

The solution for the best fit problem is now a *local* translation $\mathbf{S}_{j,k}$ on top of the rigid registration (2). To avoid computations for every single pixel position, pixels at a grid spaced Δ_{xy} are corrected in x and y position by the local translation. It is possible to consider a $3D$ map, where the height indicates the displacement either in the x or y directions — see figure 2 for an example. In order to compute the displacement for the whole set of pixels, a spline surface may be fitted to these control points, thus performing interpolation.

In this way, these displacements interpolated from a set of control points, on top of the rigid registration achieved by (15), allow computation of x^r as in (3).

Results

A case is shown on figure 3 on the next page following the major steps for the registration and a chessboard-like image assembled to show both the registration achieved and the differences in detail for the different modalities. For the mfERG/retinography registration, the global PIU value is equal to 0.111, while for the RLA/retinography registration, this value is equal to 0.097.

Discussion

In this work, two recently introduced imaging modalities for the assessment of the human retina function, from two different perspectives, i.e., the retina response

ISSN: 1727-1983

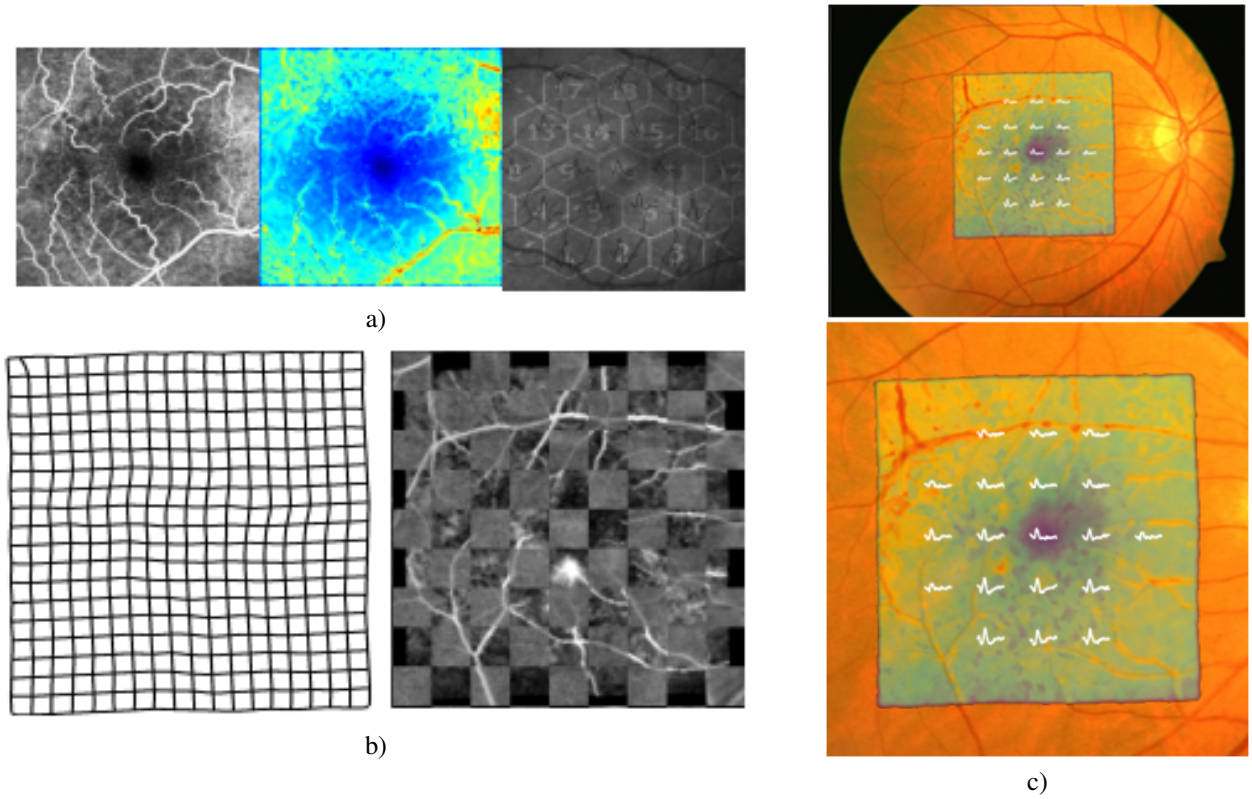

Figure 3: Modality examples and results. a) Example of image modalities: leakage analysis (fundus reference and leakage map, respectively left and center); and stimuli pattern over the multifocal electroretinogram (right). b) Grid deformation result (correcting for saccadics), where oscillations can be seen with particular incidence on specific areas (left), and a chessboard-like image assembled using interleaved blocks from both the leakage modality fundus reference and the 1st principal component of the retinography imaging modality (right). c) Final result of the registration between the multifocal electroretinogram retinal response curves (in white) and the light stimuli over the color coded leakage map (central area), represented over the background fundus photography (the retinography) which defined the coordinate system for this case. It was possible therefore to plot the curves of retina response to the position where the stimulus took place, either relatively to the distortion free leakage map or the color fundus photograph. A zoom-in (bottom) of the common area between modalities on the full-sized image (top) is presented for better visualization.

to 488 nm wavelength light stimuli and the blood-retinal barrier function, were used.

A model was established to account for the deformations due to saccadic eye movements present in imaging modalities based on single-point sensors, where fundus images are assembled from individual pixels imaged over time, and was used for registration. Moreover, to register these functional imaging modalities and simultaneously correcting for saccadics, a third imaging modality, of the morphological and not functional type, was used as ground truth so as to establish the coordinate system to which both mfERG and RLA were registered. Since these modalities are registered to a common reference, they automatically become registered to each other, as desired.

To correct for saccadics, the model herewith developed considers deformations as consisting of local translations occurring on top of a rigid transformation, assumed as having taken place priorly due to the global positioning of the eye when imaged by both equipments, on at a time. The level of success of this model in correcting for the distortions on polar coordinates due to the eventual (experienced in practice) detection of incongruent fovea center locations between modalities is crucial. Failure in doing so would prevent yielding the registration accuracy levels achieved in this work, as it would change the PIU metric values due to the poor rotation estimate resulting from important distortions on the structure of interest. Also paramount for the performance levels attained in this work was the choice of the similarity metric to be used (PIU), since we are dealing with both morphological and functional imaging modalities. In this work we also followed an approach that helped us reducing the search space, and consequently the number of parameters to be estimated in order to solve the problem at hand. This was accomplished through the assumptions made which restrict the problem's context — these are legitimate, however, since they are either in accordance with the application area of our system or based on previous experience of working with this sort of images.

ISSN: 1727-1983

Although different approaches could have been taken, we followed a path that we consider as making the best use of the knowledge and structure present on human retinal images and particularly the application area where this system will be used on.

A clear decision was made for the model of saccadics though higher order functions could have been used, as they effectively are in different application areas; it is our belief, however, that due to the impulse-like behavior of saccadics, these would have a negative effect on areas where these would have occurred. This negative effect would arise due to one of two opposite reasons:

- by not completely following displacements due to a high value for the regularization term used to smooth out the higher order polynomial;
- due to the spread of the effect of saccadics to neighborhood areas, if that regularization term is chosen on the contrary to be low enough so as to allow the system to closely follow the associated translations.

Further work consists on extending research efforts to consider even smaller FOV images of other imaging modalities eventually not containing the foveal area. In these cases, the initial estimates herewith considered cannot be applied and different methods will need to be developed and implemented.

Another improvement may be sought for the simultaneous registration of these modalities, i.e., the retinography, the mfERG and the RLA, so as to prevent propagation of registration errors. Nevertheless, given the relatively large areas covered by each of the mfERG stimuli, this was not considered fundamental for the modalities involved in our current research activities.

Conclusions

The registration of two functional imaging modalities was achieved with the simultaneous correction of the intrinsic image distortions associated to both. Therefore, the right location of the retina stimuli relatively to the blood-retinal barrier function was found. As a by-product of the necessity to correct for the intrinsic deformations present on both the mfERG and the RLA, the registration of these modalities to a retinography was undertaken. Consequently, correlation of morphological information (taken from the retinography) with functional information (given by the retina response to light stimuli from the mfERG and the blood-retinal barrier function from the RLA) was made possible, therefore allowing for the simultaneous correlation of three different modalities.

Acknowledgments

The work herewith presented was funded by the POSI/SRI/45151/2002 project, sponsored by the Fundação para a Ciência e a Tecnologia, Portugal.

References

[1] RUI BERNARDES, CONCEIÇÃO LOBO, and JOSÉ CUNHA-VAZ. Multimodal macula mapping: A new approach to study diseases of the macula. *Surv Ophthalmol*, 47(6):580–589, November–December 2002.

[2] ROBERT WEBB, G. HUGHES, and F. DELORI. Confocal scanning laser ophthalmoscope. *Appl Opt*, 26(8):1492–1499, April 1987.

[3] RUI BERNARDES, JORGE DIAS, and JOSÉ CUNHA-VAZ. Mapping the human blood-retinal barrier function. *IEEE T Bio-Med Eng*, 52(1):106–116, January 2005.

[4] FRANCE LALIBERTE and LANGIS GAGNON. Registration and fusion of retinal images—an evaluation study. *IEEE T Med Imaging*, 22(5):661–673, May 2003.

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

[6] AXEL PINZ, STEFAN BERNÖGGER, PETER DATLINGER, and ANDREAS KRUGER. Mapping the human retina. *IEEE T Med Imaging*, 17(4):606–619, August 1998.

[7] ERICH E. SUTTER and DUONG TRAN. The field topography of ERG components in man - I. The photopic luminance response. *Vision Res*, 32:433–446, 1992.

[8] MATHIAS W. SEELIGER, KRISTINA NARFSTRÖM, JENS REINHARD, EBERHART ZRENNER, and ERICH SUTTER. Continuous monitoring of the stimulated area in multifocal ERG. *Documenta Ophthalmologica*, 100:167–184, 2000.

[9] DEREK HILL and PHILIPE BATCHELOR. Registration methodology: Concepts and algorithms. In Joseph Hajnal, Derek Hill, and David Hawkes, editors, *Medical Image Registration*, The Biomedical Engineering Series, chapter 3, pages 39–70. CRC Press, 2001.

[10] CHANJIRA SINTHANAYOTHIN, JAMES F. BOYCE, HELEN L. COOK, and THOMAS H. WILLIAMSON. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *Br J Ophthalmol*, 83:902–910, 1999.

[11] MINH DO. Contourlets and sparse image expansions. In *Wavelets: Applications in Signal and Image Processing X. Proceedings of the SPIE.*, pages 560–570, 2003.

[12] PEDRO BAPTISTA, RUI BERNARDES, JOÃO FERREIRA, JORGE DIAS, and JOSÉ CUNHA-VAZ. Multimodal macula mapping: Study for rigid, perspective and non-rigid image registration. In *Proceedings of the 3rd European Medical and Biological Engineering Conference*, 2005. In press.

ISSN: 1727-1983