Robust multimodal registration of fluorescein angiography and optical coherence tomography angiography images using evolutionary algorithms

Javier Martínez-Río [a], Enrique J. Carmona [a,*], Daniel Cancelas [a], Jorge Novo [b,c], Marcos Ortega [b,c]

[a] *Department of Artificial Intelligence, Universidad Nacional de Educación a Distancia, Madrid, Spain*
[b] *Department of Computer Science and Information Technologies, University of A Coruña, A Coruña, Spain*
[c] *CITIC-Research Center of Information and Communication Technologies, University of A Coruña, A Coruña, Spain*

ARTICLE INFO

ABSTRACT

Optical coherence tomography angiography (OCTA) and fluorescein angiography (FA) are two different vascular imaging modalities widely used in clinical practice to diagnose and grade different relevant retinal pathologies. Although each of them has its advantages and disadvantages, the joint analysis of the images produced by both techniques to analyze a specific area of the retina is of increasing interest, given that they provide common and complementary visual information. However, in order to facilitate this analysis task, a previous registration of the pair of FA and OCTA images is desirable in order to superimpose their common areas and focus the gaze on the regions of interest. Normally, this task is manually carried out by the expert clinician, but it turns out to be tedious and time-consuming. Here, we present a three-stage methodology for robust multimodal registration of FA and superficial plexus OCTA images. The first one is a preprocessing stage devoted to reducing the noise and segmenting the main vessels in both types of images. The second stage uses the vessel information to do an approximate registration based on template matching. Lastly, the third stage uses an evolutionary algorithm based on differential evolution to refine the previous registration and obtain the optimal registration. The method was evaluated in a dataset with 172 pairs of FA and OCTA images, obtaining a success rate of 98.8%. The best mean execution time of the method was less than 5 s per image.

1. Introduction

Optical coherence tomography angiography (OCTA) is a noninvasive imaging modality used to evaluate the vascular flow of the retina and choroid. This technique creates a blood flow image by comparing the differences between repeated OCT cross-sections of a given location in the retina: signal changes associated with movements of red blood cells are larger than the ones related to static tissue [1]. OCTA is widely used clinically to diagnose and grade different types of ophthalmological diseases, such as retinal artery/vein occlusions, diabetic retinopathy (DR), and age-related macular degeneration (AMD), among others. On the other hand, fluorescein angiography (FA) and indocyanine green angiography (ICGA) are invasive imaging modalities in which dye is injected into the bloodstream to highlight the blood vessels in the outermost layer of the retina. They are useful for the study of choroidal and retinal neovascularization, helping in the diagnosis and grading of relevant diseases such as DR and AMD.

These three imaging modalities have been compared in the related literature [1–3]. For example, the network of vessels visualized using FA/ICGA only corresponds to the superficial capillary plexus, while OCTA provides improved visualization of all the vascular layers, including different depths of the retinal layer (superficial and deep capillary plexuses). In addition, and unlike OCTA, FA/ICGA are invasive modalities, relatively expensive, and time-consuming. On the contrary, FA/ICGA are less prone to artifact and provide images with a wider field of view than OCTA. For example, the automated scan protocols that are typically available in clinical practice for OCTA images are 3×3 mm and 6×6 mm.

At present, there is a growing interest in using a multimodal approach where these imaging modalities can help to improve the identification of pathological changes, given that the blood flow information provided by FA/ICGA and OCTA is complementary [2]. However, to facilitate the comparison of this information, it is necessary to apply a previous registration process that allows the expert clinician to

identify the common area covered by the OCTA image in the FA or ICGA image. Considering that multimodal registration refers to the process of using two or more images of different modalities in spatial alignment, this task has normally been approached manually or semi-automatically [4–9], although recently it has already started to be addressed automatically [10].

In relation to the manual or semi-automatic methods, we briefly describe some of the most recent and relevant ones. For example, in Ref. [9], the OCTA images were manually aligned with the FA/ICGA images using the inner retinal vasculature as a reference. Then, all the visible microaneurysms (MA) in the registered regions of each imaging modality were manually annotated and, subsequently, the counted numbers were compared to evaluate the degree of agreement. In Ref. [7], FA and OCTA images were manually registered using retinal vascular landmarks and, next, each FA image was finally cropped to the corresponding OCTA dimensions (3×3 mm or 6×6 mm). The objective was to facilitate, within the registered regions, quantitative analysis of the number of MAs and foveal avascular zone (FAZ) area in patients with diabetic macular edema. In Ref. [6], OCTA and FA/ICGA images were aligned by using a semiautomatic registration tool that translates, rotates, and rescales the OCTA image to accurately match three user-specified landmark locations on the respective images. Then, two observers manually annotated signs of abnormal choroidal blood flow in the registered regions and the overlap of these annotations was calculated. The final objective was to evaluate the degree of consensus in the annotations that were made in each type of image.

On the other way, the fully automatic approach utilized in Ref. [10] consisted of the following steps. First, a deep learning algorithm to segment blood vessels and detect vessel junctions in FA/ICGA and OCTA images was used. Then, the registration process was done in two steps. In the first one, the vessel junctions were used to estimate the scaling, translation, and rotation needed to do a first rough registration. Next, in the second step, automated fine-tuning was applied using vessels, the original images, and an image registration toolbox called elastix [11]. Finally, different choroidal neovascularization lesion parameters and FAZ size were compared using OCTA-ICGA and OCTA-FA, respectively.

Presented the current interest in the ophthalmology community about taking advantage of the different information provided by the different imaging modalities above described, it is desirable to address the problem of automatically registering the images produced by these techniques in order to align and extract the common regions to be compared. In this work, we propose a three-step approach for robust multimodal registration of FA and superficial plexus OCTA images. In the first step, a preprocessing stage is applied to the original FA and OCTA images to segment their main vessels. Then, a second step based on template matching uses the vessel information to obtain a first rough registration, which is subsequently refined by means of an evolutionary algorithm in a third step. Note that the problem of registering FA and deep plexus OCTA images is not addressed here. It is more complex, given that the network of vessels and the rest of information shown in FA images belong to the superficial plexus.

The following should also be considered: The number of cross-sectional OCT scans is limited by the scanning speed of the instrument and, as a result, a larger field of view will have reduced density and resolution. This means that a 3×3 OCTA image has a higher level of

detail per mm^2 than a 6×6 OCTA image, but the field of view of the former is smaller than the latter. That is why, depending on the size of the retinal area to analyze and the type of pathology to study, the expert can be interested in using a 3×3 or 6×6 OCTA image to be registered with the FA image. For example, avascular areas can be detected more accurately on 3×3 images than on 6×6 images. In addition, since the 3×3 and 6×6 OCTA images are captured separately, the registration between them is unknown *a priori*. Therefore, the FA and 3×3 OCTA registration result cannot be used to register FA and 6×6 OCTA images, or vice versa. Consequently, the FA and 3×3 OCTA image registration problem is different from the FA and 6×6 OCTA image registration problem. In this context, the proposed method will address both problems.

Although other metaheuristic-based approaches have been proposed in the related literature to address the multimodal medical image registration problem [12–15], the main contribution of this work is twofold. First, we show evidence on how an evolutionary algorithm based on differential evolution can be used successfully in the multimodal image registration problem, specifically for the case of FA and OCTA images. The novel idea consists of evolutionarily searching for the parameters that define an affine transformation that produces optimal (or near-optimal) registration between both types of multimodal images. Each of these parameters is directly associated with each of the linear transformations (translation, scaling, rotation and shear) involved in the global affine transformation. This favors a more intuitive adjustment of the definition domain of each parameter before performing the search process. Second, we also present a procedure to automatically evaluate the success or failure of the registration made by our method.

The rest of the manuscript is organized as follows: Section 2 describes the features of the image dataset used in the different experiments and also introduces differential evolution. Section 3 explains in detail the approach that was proposed for registering FA and OCTA images. Section 4 shows and analyzes the results obtained in the experiments performed to validate our approach. Finally, Section 5 exposes the main conclusions of this work.

2. Materials and background

The "Comité de Ética da Investigación de Santiago-Lugo" committee, belonging to the "Rede Galega de Comités de Ética da Investigación" attached to the regional government "Secretaría Xeral Técnica da Consellería de Sanidade da Xunta de Galici", approved this study, which was conducted in accordance with the tenets of the Helsinki Declaration. This study was carried out retrospectively on existing data that have previously been anonymized. The validation process was done using a dataset that contains a total of 86 cases: 31 healthy cases and 55 pathological cases, the latter showing signs of retinal vein occlusion. In turn, each case corresponds to a patient's eye and consists of one FA image and two superficial capillary plexus OCTA images, each one with a different level of zoom: 6×6 *mm* and 3×3 *mm*. Therefore, a total of 172 registrations are possible (86 pairs of FA and OCTA$_{3\times3}$ images plus 86 pairs of FA and OCTA$_{6\times6}$ images). The image resolutions are 1536×1536 and 320×320 pixels for FA and OCTA, respectively, but the 3×3 and 6×6 OCTA images have the same resolution. The FA and OCTA images were taken using the Optical Coherence Tomography capture device *DRI OCT Triton - Topcon Corp*.

The data set contains information on a total of 29 patients, whose ages range from 46 to 92 years. A number of cases, ranging from two to four, may correspond to the same patient. The different cases in a patient may refer to the type of eye (left or right), or to the same type of eye but at different instants of time.

The OCTA images are normally centered in the macular area and some of them present motion artifacts that are typical of this kind of imaging modality in their routine capture process. They are produced by involuntary eye motions, such as microsaccades, that occur during the 3–5 s that normally takes a single OCTA volumetric [16]. These artifacts

Table 1
Distribution of the different pairs of FA and OCTA images depending on the diagnosis of the patient's eye (healthy or pathological), the zoom level of the OCTA image (3×3 or 6×6), and the presence or absence of motion artifacts in the OCTA image.

	3×3 OCTA		6×6 OCTA	
	no-artifacts	artifacts	no-artifacts	artifacts
Healthy	23	8	25	6
Pathological	48	7	48	7

are clearly visible as horizontal or vertical white lines that disrupt the continuity of vascular networks. The existence of these characteristic motion artifacts reinforces the idea of providing accurate and robust registration methodologies that may adequately handle these types of situations, which are common in clinical practice when working with OCTA images. Table 1 summarizes the number of FA and OCTA image pairs in the dataset for the different categories described above. Fig. 1 shows two examples of healthy and pathological eye cases, illustrating different scenarios in the used dataset.

Finally, by way of introduction, we present the evolutionary paradigm called differential evolution (DE) [17,18], which will form part of the registration method described in the next section. DE is an evolutionary algorithm (EA) that can optimize non-differentiable, nonlinear and multimodal cost functions by iteratively improving a set of n-dimensional real-valued vectors. Basically, DE follows the typical stages of an EA (population initialization, parent selection, mutation, recombination, and survival selection), but it is characterized by a particular implementation of the mutation operator. Although there are different DE variants [18], we will use the one denoted by "DE/best/1/bin" (see Algorithm 1), which works as follows. In each generation and for each individual of the population, $\mathbf{x}_i$, the mutation operator produces a new mutated individual, $\mathbf{m}_i$, given by Eq. (1), where $F \in [0, 2]$ is a constant called the *differential weight* and chosen by the user, $\mathbf{x}_{best}$ is the best individual in the current generation, and $\mathbf{x}_p$ and $\mathbf{x}_q$ are two individuals used to build the called *difference vector*; both individuals are selected from the current population, at random, without replacement, and different from $\mathbf{x}_i$.

$$\mathbf{m}_i = \mathbf{x}_{best} + F \cdot (\mathbf{x}_p - \mathbf{x}_q) \tag{1}$$

To create a new child, the recombination operator selects each of its components from $\mathbf{x}_i$ or $\mathbf{m_i}$ based on CR, where $CR \in [0, 1]$ is called the *crossover probability* and is also defined by the user. Finally, those children $\mathbf{y}_i$ that are better than their respective parents $\mathbf{x}_i$ replace the latter in the new population. The notation "DE/x/y/z" is interpreted as follows: label x refers to how the vector that is added to the weighted difference vector is selected (in our case, the best of the current generation is used, i.e., "$x = best$"); label y denotes how many couples of individuals are involved in the computation of the difference vector (in our case, "$y = 1$"); and label z refers to the type of recombination used (in our case, binomial, i.e., "$z = bin$", such as is shown in lines $8 - 15$ of Algorithm 1).

3. Method

The proposed methodology for image registration of FA and OCTA images consists of three main stages, which are illustrated in Fig. 2. First, a preprocessing stage is applied to both OCTA and FA images in order to segment the main vessels. Next, a rough registration is done applying Template Matching based on cross-correlation. Finally, a fine registration is made using an EA, the objective of which is to find the best affine transformation that, taking advantage of the rough registration obtained in the previous stage, allows the OCTA image to be optimally registered on the FA image.
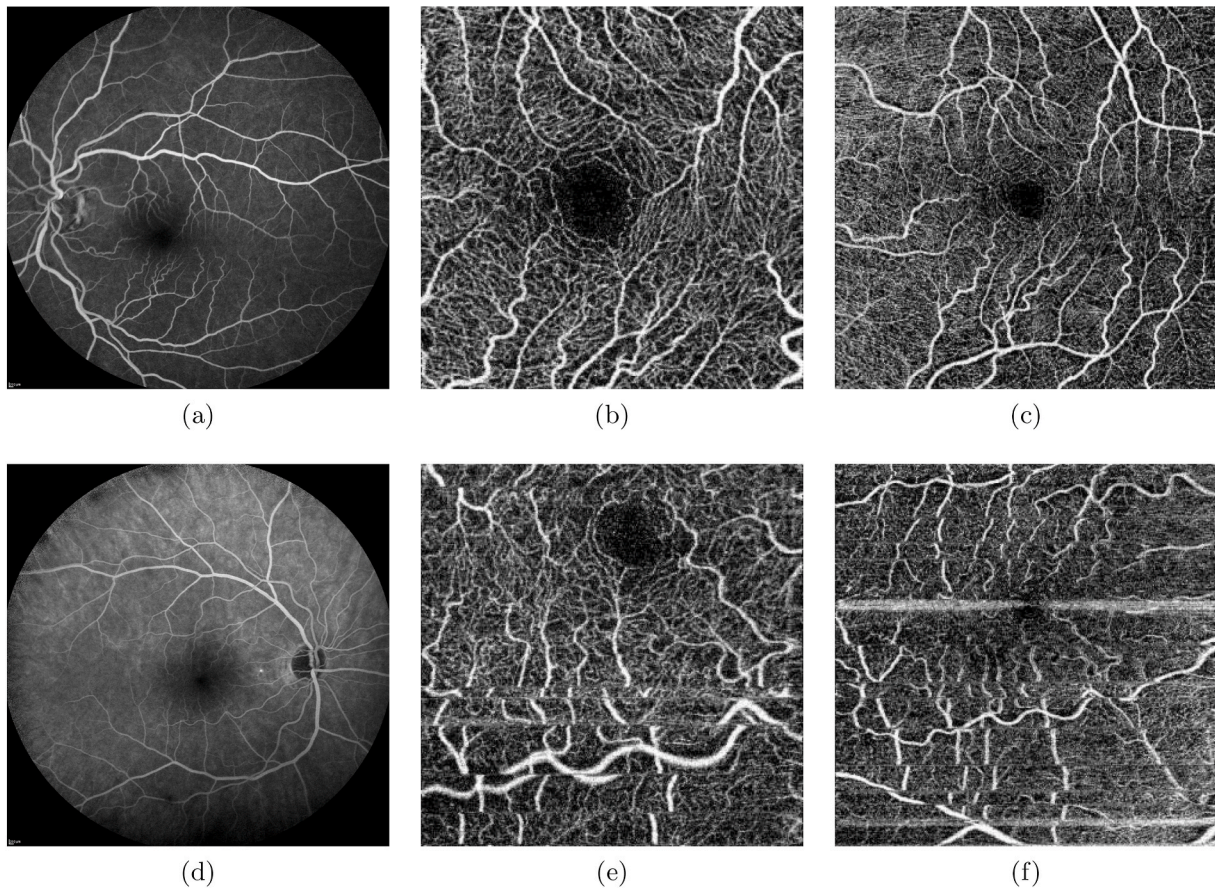

(a) (b) (c)

(d) (e) (f)

Fig. 1. Examples of healthy and pathological eye cases in the first and second row, respectively: (a) FA image; (b) 3×3 *mm* superficial plexus OCTA image without artifacts; (c) 6×6 *mm* superficial plexus OCTA image without artifacts; (d) FA image; (e) 3×3 *mm* superficial plexus OCTA image with motion artifacts; (f) 6×6 *mm* superficial plexus OCTA image with motion artifacts.

Algorithm 1 Pseudo-code of the DE algorithm ("DE/best/1/bin" variant), where a maximization problem is assumed.

Inputs:

$f : \mathbb{R}^n \to \mathbb{R}$, fitness function to maximize

$CR \in [0,1]$, crossover probability

$F \in [0,2]$, differential weight

S, population size

Output:

the best individual from final population

```
(01) Initialize P ← {x_1, ..., x_S}
```
(02) **until** (termination-criterion)

(03) $\mathbf{x}_{best} \leftarrow \mathbf{x}^* \mid \mathbf{x}^* \in \mathbf{P} \wedge \forall \mathbf{x} \in \mathbf{P} : f(\mathbf{x}^*) \geq f(\mathbf{x})$

(04) **for** each $i \in \{1, ..., S\}$

(05)   Pick-Randomly $\{p,q\} \in \{1,..,S\}$ with $i \neq p \neq q$

(06)   Pick-Randomly $\beta \in \{1, ..., n\}$

(07) $\mathbf{m}_i \leftarrow \mathbf{x}_{best} + F \cdot (\mathbf{x}_p - \mathbf{x}_q)$

(08)   **for** each $j \in \{1, ..., n\}$ *%Binomial recombination*

(09) $r_j \equiv U(0,1)$

(10)   **if** $r_j < CR$ or $j = \beta$

(11) $y_j \leftarrow m_{i,j}$

(12)   **else**

(13) $y_j \leftarrow x_{i,j}$

(14)   **end-if**

(15)   **end-for**

(16) $\mathbf{y}_i \leftarrow \mathbf{y}$

(17) **end-for**

(18) **for** each $i \in \{1,..,S\}$ *%Generational replacement*

(19)   **if** $f(\mathbf{y_i}) > f(\mathbf{x}_i)$ **then** $\mathbf{x}_i \leftarrow \mathbf{y_i}$
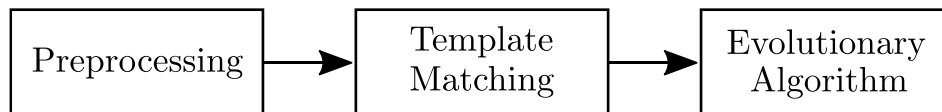
(20) **end-for**

(21) **end-until**

Fig. 2. Main stages of the proposed methodology.

3.1. Preprocessing

The final goal of the preprocessing stage is to detect and isolate the main retinal vessels in each pair of OCTA and FA images, given that these anatomical structures represent the main source of common information between both types of images. This stage consists of several steps that are applied in a sequential pipeline: normalization, elimination of noise, and main vessel segmentation. Note that this pipeline is applied to both types of images (OCTA and FA).

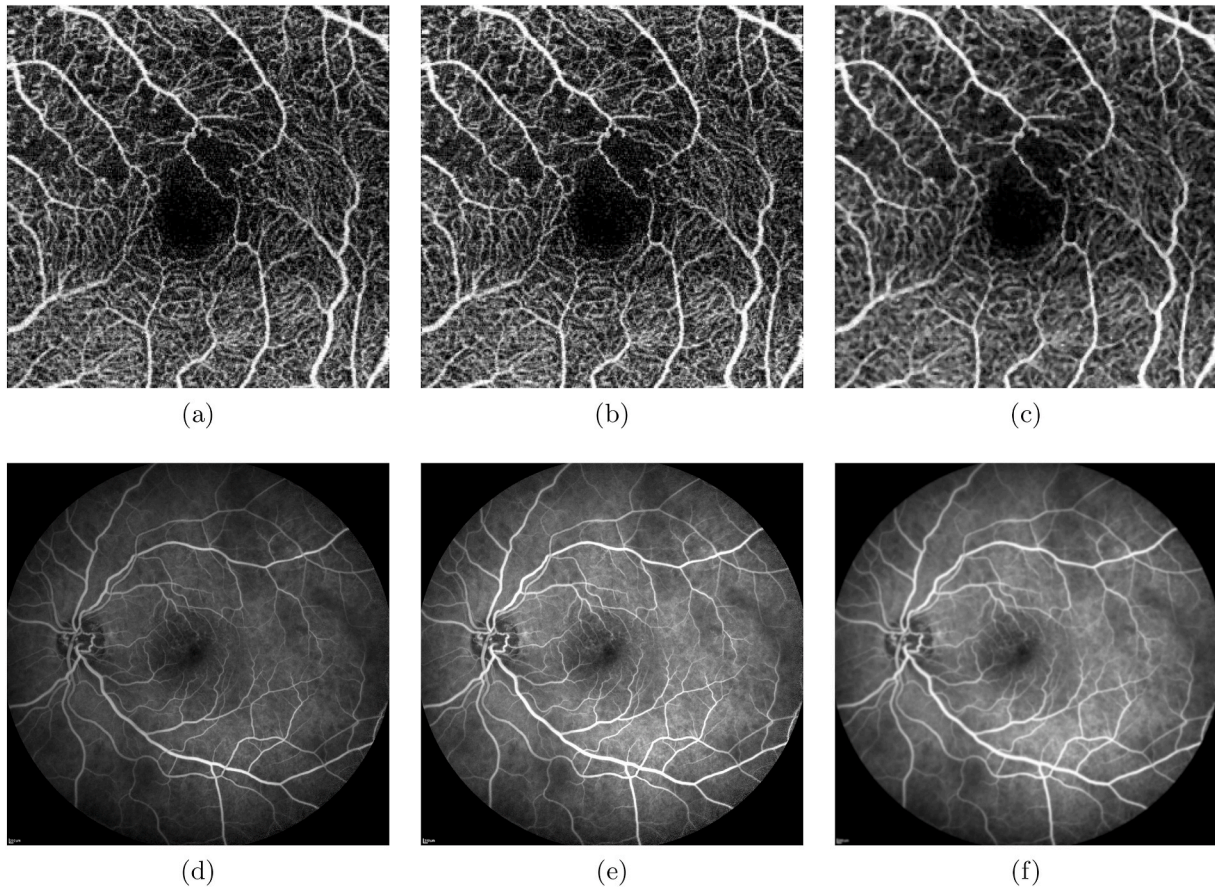#### 3.1.1. Normalization and denoising

The normalization process is done to increase the contrast of each image, thus covering the entire range of gray levels [0, 255]. However, to avoid the adverse effect of outliers in this process, the bottom 1% and top 1% of the total pixel number with lower and higher values in the image being processed are saturated to 0 and 255, respectively. Next, we apply a filter of size $m \times m$ pixels to reduce noise in the images. Since the

imaging modalities are different, we also experimentally verified that the applied filters also behaved differently. We obtained the best results using a Gaussian filter on the FA images and a median filter on the OCTA images. For choosing the size of each filter, the resolution of each type of image was taken into account. Thus, the size was large enough to eliminate as much noise as possible, but small enough to preserve the quality of the main vessels. Finally, we selected $m = 19$ and $m = 3$ for the size of the Gaussian and median filters, respectively. Fig. 3 shows output examples of applying each one of these steps in both types of images.

3.1.2. Main vessel segmentation

As indicated, both imaging modalities specially highlight the blood flow and the vascularity. Therefore, we assume that the extraction of the main vessels is adequate to support the posterior registration process and provide a suitable and robust method.

The main vessel segmentation process consists of several steps that are sequentially applied, taking as the starting image the one produced by the denoising process (see previous section). The first step allows us to enhance the vessels by means of a top-hat morphological operator with a disk-shaped structuring element (SE). The radius of the SE

Fig. 3. Examples of applying the normalization and denoising steps: (a) and (d) original OCTA and FA images, respectively; (b) and (e) outputs of applying the normalization step to (a) and (b), respectively; (c) and (f) results after applying the denoising step to (c) and (f), respectively.

depends on the resolution of each type of image in our dataset: radii equal to 31 and 15 were respectively chosen to enhance the vessels in the FA and OCTA images.

The next step is to create a binary mask, where pixels equal to one represent the main vessels. The method implemented for accomplishing this task combines Hessian-based filtering and thresholding based on hysteresis filter [19]. The Hessian matrix is computed by convolving the image with the second derivatives of the Gaussian kernel in both directions (rows and columns). The only parameter to configure this filter depends on the standard deviation σ used for the Gaussian kernel, which is used as a weighting function for the auto-correlation matrix. It was experimentally tuned to $\sigma = 3.5$ and $\sigma = 4$ pixels for the FA and OCTA images, respectively.

Next, the idea with the hysteresis filter is to start from a maximum threshold (Th_{max}) that guarantees that the image pixels that exceed this limit correspond to vessels. Then, a new threshold is chosen (Th_{tmp}), which is progressively decreased in each iteration from Th_{max} to a minimum threshold (Th_{min}). So, in each iteration, those image areas containing pixel intensities that are above the Th_{tmp} will be considered fragment of vessels only if they are connected to image areas that were previously labeled as vessels. Thus, the Th_{tmp} determines which new areas of the image can be seen as possible continuations of those other areas that were already labeled as vessels with high confidence. The election of the Th_{max} value is direct and intuitive. However, in order to facilitate the right tuning of the Th_{min}, we make its value dependent on the desired percentage of active information (ratio between mask pixels with a value equal to 1 and total pixels). Thus, the Th_{max} was set to 200 and 220 for the FA and OCTA images, respectively, and, regarding the Th_{min}, 20%, 15%, and 30% were the percentages chosen for the FA, 3 × 3 OCTA, and 6 × 6 OCTA images, respectively. All these threshold values

were obtained experimentally by trial and error. However, here it should be considered that the resolutions of 3 × 3 and 6 × 6 OCTA images are the same in our dataset and the average ratio of vessel pixels between both types of images, which was roughly calculated, is 1.6. Therefore, it makes sense to consider $Th_{min_{6 \times 6}}$ to be twice $Th_{min_{3 \times 3}}$.

The obtained mask in the previous step tends to overestimate the thickness of the vessels. Then, each mask is applied to the respective output image of the top-hat operator and the result is a new image that contains the main vessels in gray levels. In this way, the actual information of the vessels is retrieved. From now on, the OCTA and FA images obtained after this preprocessing stage will be denoted by OCTA vessel (OCTAv) and FA vessel (FAv) images, respectively. Fig. 4 shows an example of the different steps followed to segment the main vessels.

Finally, Table 2 summarizes the set of parameter-value pairs associated with the preprocessing stage. Note that this configuration is dependent on the resolution of the FA and OCTA images belonging to the dataset used to validate our method (see Sec. 2). Therefore, the use of the other resolutions will involve an additional step dedicated to adequately tune the value of these preprocessing parameters.

3.2. Template matching

Since the OCTA and FA images are always taken with the patient's eye in similar orientations, small differences in the rotation angle are expected. Therefore, we assume that a rough transformation for registering both images can be based on scaling and translation. However, the scale factor can be easily determined by the information provided by the device used to obtain both types of images. Consequently, the transformation to be applied will only require one translation. However, instead of directly calculating this transformation, we propose an
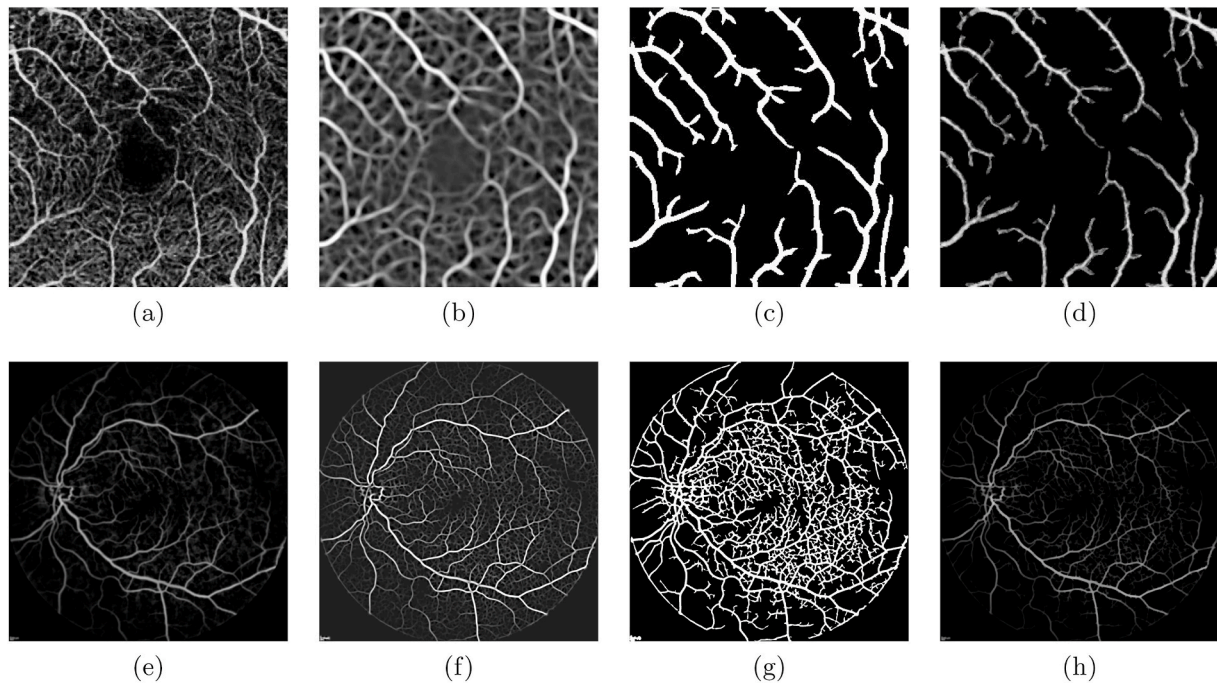
Fig. 4. Examples of the main vessel segmentation step: (a) and (e) results of applying, respectively, the top-hat operator to the images (c) and (f) in Fig. 3; (b) and (f) outputs after using the Hessian matrix to (a) and (e), respectively; (c) and (g) binary masks obtained after applying the hysteresis filter in (b) and (f), respectively; and (d) and (h) main vessels (in gray level) obtained after applying the binary mask to the images (a) and (e), respectively.

exhaustive search based on template matching to obtain a first approximation to the problem of registering OCTA and FA images.

In general, template matching (TM) applied to image-processing consists of finding the region, inside an image *frame,* with which an image *template* has maximum similarity. Usually, the template is completely contained in the frame and the size of the former is quite smaller than the size of the latter. From an algorithmic point of view, TM involves the following steps. First, the template is shifted, one pixel at a time, from left to right and top to bottom over the frame. Second, at each location, the similarity between the template and the frame subwindow encompassed by the template is calculated. Finally, the frame region, where the similarity with the template is maximum, is provided as the process output. In our case, the OCTAv and FAv images are used as the template and frame, respectively. Then, the TM output will provide the best region of the FA image with which the OCTA image has maximum similarity. In the literature related to image processing, various numerical measures have been used to calculate similarity, such as the mean absolute error, root mean square error, and correlation, among others. However, the correlation, in addition to its reliability and accuracy [20], has two interesting properties: it is independent of any offset or linear transformation in the set of pixel values to match and its value will always belong to the interval $[-1, 1]$ if its normalized version is used.

In our proposal, the angiography image maintains its original size, while the OCTA image is accordingly scaled to assure that the final scale in both images is the same. Without loss of generality, the opposite could
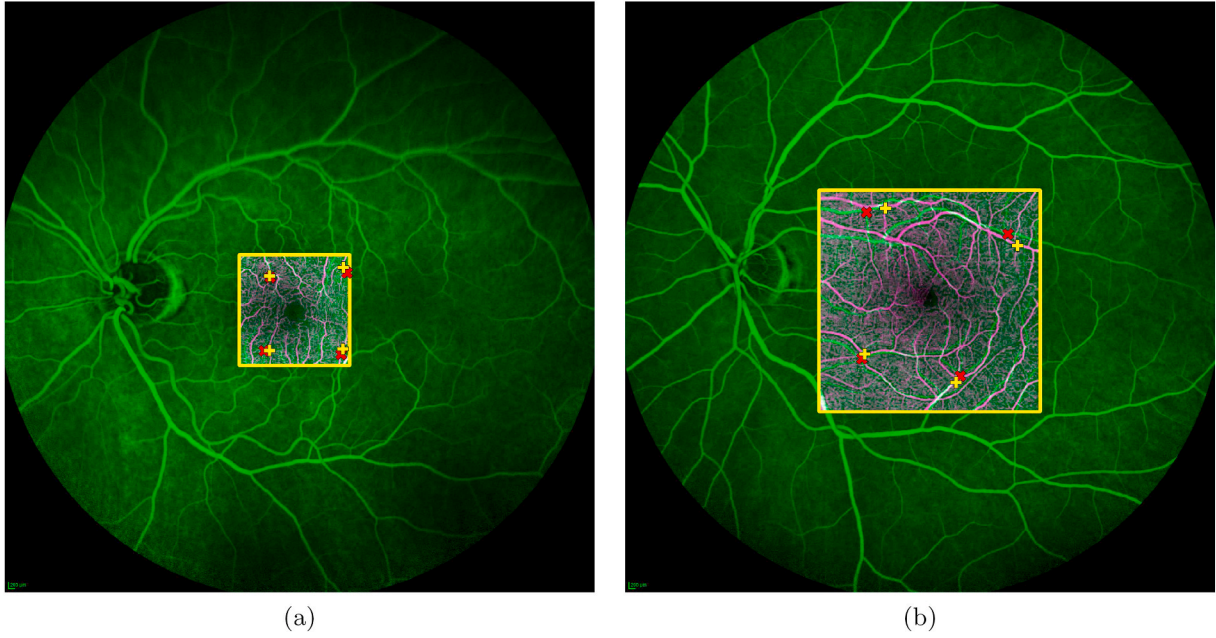
also have been done: OCTA images retain their size and FA images are rescaled. The scale factors to be applied along the *x* and *y* axes are the same, $s_{x_0} = s_{y_0} = s_0$, in such a way that s_0 for the 3×3 and 6×6 OCTA images is 0.9 and 1.8, respectively. Note that the original size (in pixels) of the 3×3 and 6×6 OCTA images in our dataset is the same, but the actual size (in mm) of the side of the square region covered by a 3×3 OCTA image is half the side size of the square region covered by a 6×6 OCTA image. Therefore, this explains why the scale to be applied in each direction (vertical and horizontal) is the same (s_0) and, for the 3×3 OCTA images, it is half that the one applied in the 6×6 OCTA images.

It is important to highlight that TM is applied to the images obtained after applying the main vessel segmentation stage (see Fig. 4 (d) and (h)). Here, it must be remembered that both imaging modalities highlight the blood flow and the vascularity and, therefore, the rest of non-vessel information, in addition to being different in both types of images, can provide distractions that make the matching process to fail. Consequently, we only work with the main vessel information (OCTAv and FAv images). Finally, as a measure of similarity, the so-called zero normalized cross correlation (ZNCC) is used [21]. Specifically, the ZNCC of a $p \times q$-size template, $T(x, y)$, with a $p \times q$-size frame subwindow, $F_s(x, y)$, is:

$$\text{ZNCC}(T, F_s) = \frac{1}{n} \sum_{x,y} \frac{1}{\sigma_{F_s} \sigma_T} (F_s(x, y) - \mu_{F_s})(T(x, y) - \mu_T) \qquad (2)$$

where $n = p \times q$, μ_T and μ_{F_s} are the average of $T(x, y)$ and $F_s(x, y)$, respectively, and σ_T and σ_{F_s} are their respective standard deviations. It should be stressed that $F_s(x, y)$ represents that region of the frame, $F(x, y)$, encompassed by $T(x, y)$ when this is located in a determined position on $F(x, y)$. Therefore, $F_s(x, y)$ will vary each time $T(x, y)$ moves a position from left to right and top to bottom over $F(x, y)$.

In order to illustrate the TM output, Fig. 5 shows a pair of examples of registration. In the overlap zone (yellow square), the magenta and green colors represent vessels in OCTA and FA images, respectively, and white appears when the vessels in both images overlap. Despite the fact that this stage provides an acceptable approximation, we can appreciate that the obtained registrations are not perfect. This degree of discrepancy can

Table 2
Final parameter configuration in the preprocessing stage.

Operator	Parameter	FA	3×3 OCTA	6×6 OCTA
Median filter	size (px)	–	3×3	3×3
Gaussian filter	size (px)	19×19	–	–
Top-Hat	SE radius (px)	31	15	15
Hessian filter	σ	3.5	4.0	4.0
Hysteresis filter	$Th_{\max}$	200	220	220
	$Th_{\min}$	20%	15%	30%

Fig. 5. Examples of TM-based registration: (a) 3 × 3 OCTA image; (b) 6 × 6 OCTA image (see Sec. 3.2 for more details). Each pair of corresponding points is represented by the symbols + and ×.

also be intuitively assessed by looking at the distance between each pair of corresponding points, which are shown in each pair of registered images (see Section 4 for how these pairs of points were obtained). Consequently, a finer registration process is desirable, which is explained in the following section.

3.3. Evolutionary algorithm

Using the rough registration obtained by TM as a starting point, an EA is used to optimize the parameters of a geometric transformation that, applied to the OCTAv image, allows us to obtain an optimal (or close-optimal) registration on the FAv image. Here, we assume that an affine transformation is sufficiently good to correct the deformation of projecting the retinal concave surface into a plane. Additionally, the optimality criterion is based on the degree of correlation obtained between the transformed OCTAv image and its corresponding FAv image, in such a way that we also assume that the higher the correlation value, the better the quality of the registration.

A 2D affine transformation, which transforms the point (x, y) into the transformed point (x', y'), is defined by the following matrix:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3}$$

with:

$$\begin{aligned} a_{11} &= s_x \cdot \cos(\theta_r) \\ a_{12} &= -s_y \cdot \sin(\theta_r + \alpha_s) \\ a_{21} &= s_x \cdot \sin(\theta_r) \\ a_{22} &= s_y \cdot \cos(\theta_r + \alpha_s) \end{aligned} \tag{4}$$

and where t_x and t_y specify the displacements along the x and y axes, respectively, s_x and s_y represent the scale factors along the x and y axes, respectively, and θ_r and α_s specify the rotation and shear angles, respectively. Therefore, the affine transformation is defined by a six-dimensional real-valued vector:

$$\mathbf{z} = (t_x, t_y, s_x, s_y, \theta_r, \alpha_s) \tag{5}$$

Finally, we can formalize the optimization problem to be solved by the

EA:

$$\underset{\mathbf{z} \in \mathbb{R}^6}{\arg \max} \ \ ZNCC(\Psi(OCTAv, \mathbf{z}), FAv) \tag{6}$$

where $ZNCC(*,*)$ is given by Eq. (2), and $\Psi(X, \mathbf{z})$ is an affine transformation (see Eq. (3)), which is defined by the vector of parameters $\mathbf{z}$ (see Eq. (5)), and is applied to X image (centered at the origin of coordinates).

3.3.1. Individual representation and fitness function

In order to solve the optimization problem presented in Eq. (6) by means of an EA, we need to establish a way to encode and evaluate each potential solution (chromosome) in the population. Fortunately, the codification is direct: it will be enough to use a 6-dimensional real-valued vector, see Eq. (5), to encode the parameters that define the desired affine transformation. However, to take advantage of the rough registration obtained by the TM stage, we applied two strategies. In the first one, the individual that defined the TM-based affine transformation solution, $(t_{x_0}, t_{y_0}, 1, 1, 0, 0)$, was included in the initial population, being (t_{x_0}, t_{y_0}) the translation obtained by TM in the previous stage (see Sec. 3.2). In the second strategy, for each gene (parameter) on the chromosome, we established a definition domain based on the neighborhood of its respective TM-based affine transformation parameter; in this way, we reduced the search space of the EA, increasing the possibility of finding the optimal affine transformation. The upper and lower extremes of the definition interval for each parameter were chosen in the following way: $t_{x_0} \pm 30$ and $t_{y_0} \pm 30$ pixels for t_x and t_y, respectively; $1 \pm 5\%$ of w_0 and h_0 for s_x and s_y, respectively, being $w_0 \times h_0$ the original template size; and $\pm 15°$ and $\pm 5°$ for θ_r and α_s, respectively. These ranges of variation were experimentally chosen, but they were large enough to allow a flexible registration, and small enough to take advantage of the rough registration obtained by the TM stage. In any case, the criteria finally chosen to delimit each domain of definition are not critical and similar ones could have been taken.

Finally, in relation how to evaluate each individual, the solution is also direct. It will be enough to use the function to maximize in Eq. (6) as the fitness function, that is:

$$fitness = ZNCC(\Psi(OCTAv, \mathbf{z}), FAv) \tag{7}$$

Table 3
Configuration of parameters belonging to the EA based on DE.

DE Parameters	Value
Variant	*best/1/bin*
Population size (S)	20 + 1
Initialization method	LHS
Max. number of generations	100
Differential weight (F)	0.50
Crossover probability (CR)	0.75

3.3.2. DE for multimodal image registration

We have selected DE to play the role of EA in our method for several reasons. Firstly, DE is simple and straightforward to implement. Secondly, the number of hyperparameters to tune in DE is lower than other competitive real parameter optimizers [22]. Third, in a natural way, DE manages solutions represented by fixed-dimensional real-valued vectors. Note that the potential solutions in our problem are represented by 6-dimensional real-valued vectors. Finally, DE has been successfully used for image segmentation in different medical fields [23–25]. However, considering the representation used and the characteristics of the problem to be solved, other EA candidates could be used, such as evolution strategies [26], genetic algorithm [27] or covariance matrix adaptation evolution strategy [28] or even other meta-heuristics, such as particle swarm optimization [29], among others. In any case, it is important to note that the method proposed here is independent of the method used to optimize the six parameters that define the target affine transformation.

In order to ensure that the initial population is a sample representative of the search space, an intelligent initialization was chosen. Specifically, 20 individuals were near-randomly obtained using the Latin hypercube sample (LHS) method [30]. An extra individual, which represents the solution of the TM-based affine transformation, was also added to the initial population (see Sec. 3.3.1). Given that DE implicitly implements elitism, the registration obtained by DE will always be equal to or better than the rough registration obtained by TM, ensuring that in the worst case, this adequate solution is provided by the method.

The execution of the EA stops when the maximum number of generations is reached or each population individual is at a Euclidean distance equal to or less than 10^{-2} from the current best individual. The values of the rest of DE hyperparameters are collected in Table 3. In particular, the values for the *F*, *CR* and maximum number of generations were experimentally tuned. Lastly, in relation to the DE variant finally used, we experimentally verify that the "DE/best/1/bin" variant outperformed the classical "DE/rand/1/bin" variant. The improvement mainly resulted in a shorter time to reach the solution. Thus, the inclusion in the initial population of a solution close to the optimal solution (TM-based registration) seems to be better exploited by the DE variant used than by the classical DE variant.

By way of illustration, Fig. 6 shows the results of applying DE-based registration to each of the two results obtained by TM in Fig. 5. Now, we can appreciate that the obtained registrations are notably improved. Note that, in the overlap zone (yellow square), white color appears when the vessels in both images overlap. In addition, the position of the four pairs of corresponding points used as the ground truth in each pair of registered images is also shown (see Section 4 for more details). We can see that the distances between these pairs of points have decreased when compared with the respective distances obtained in the TM-based registration (see Fig. 5).

4. Results and discussion

In this section, we conduct different experiments to evaluate and validate the proposed robust registration method by using the image dataset described in Sec. 2. It is important to highlight that all the results that are shown below were obtained with the same parameter settings (see Tables 2 and 3). We also analyze the impact of both the pre-processing parameter configuration and the computational cost on the performance of the method. Finally, we describe a new procedure to automatically evaluate the success or fail of each registration obtained.

4.1. Method evaluation

We performed a first experiment where we check if the DE-based registration really represents an improvement over the TM-based registration. Thus, we computed the ZNCC value for the registration obtained in each pair of FA and OCTA images at the end of the TM and DE stages and, finally, we calculated their mean values, $\overline{ZNCC}_{TM}$ and

(a)   (b)
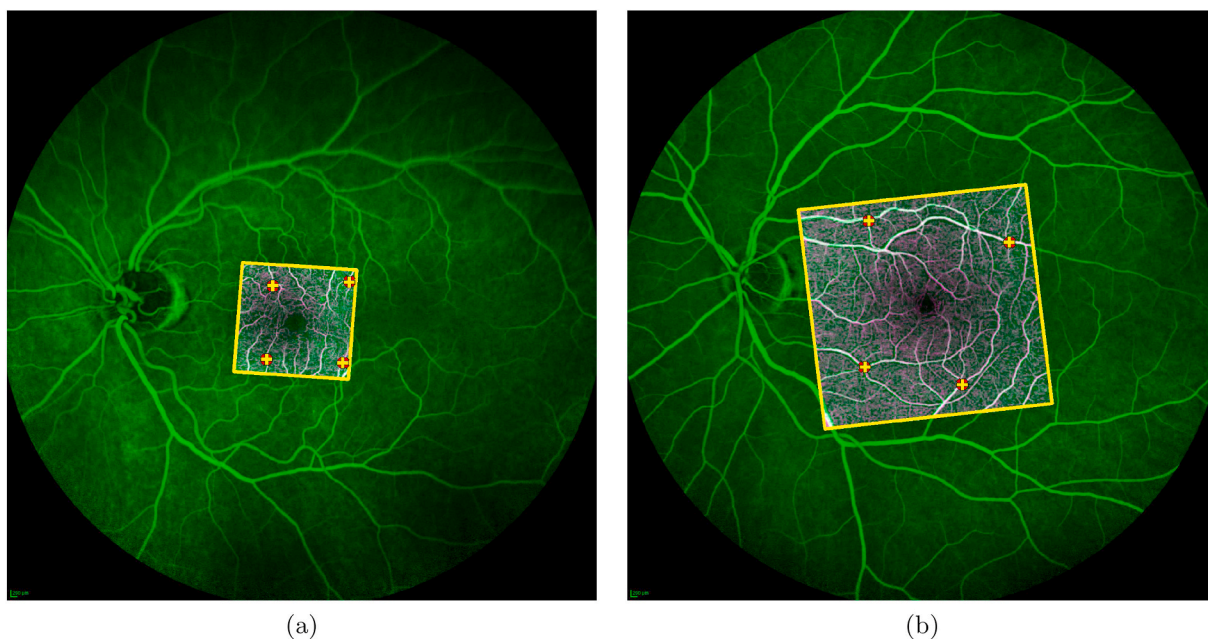
Fig. 6. Examples of DE-based registration: (a) 3 × 3 OCTA image; (b) 6 × 6 OCTA image. The improvement is evident when these results are compared with those obtained using TM-based registration (see Fig. 5). Each pair of corresponding points is represented by the symbols + and ×.

Table 4

Result of averaging the ZNCC values at the end of the TM- and DE-based registration stages, according to the proposed method. The standard deviation is represented by SD.

	$\overline{\text{ZNCC}}_{\text{TM}}$	mean($\overline{\text{ZNCC}}_{\text{DE}}$) $\pm$ SD
3×3 OCTA	0.363	0.549 ± 0.001
6×6 OCTA	0.350	0.624 ± 0.002

Table 5

Result of averaging the ZNCC values at the end of the TM- and DE-based registration stages, but re-configuring these stages to simultaneously considering both the vessel and non-vessel information. The standard deviation is represented by SD.

	$\overline{\text{ZNCC}}_{\text{TM}}$	mean($\overline{\text{ZNCC}}_{\text{DE}}$) $\pm$ SD
3×3 OCTA	0.307	0.450 ± 0.001
6×6 OCTA	0.288	0.467 ± 0.001

$\overline{\text{ZNCC}}_{\text{DE}}$, respectively. Note that EAs are stochastic methods, therefore, we run ten times the proposed registration method, calculating the mean, mean($\overline{\text{ZNCC}}_{\text{DE}}$), and standard deviation (SD) of the ten $\overline{\text{ZNCC}}_{\text{DE}}$ values. Note that TM is a deterministic method and, therefore, the same $\overline{\text{ZNCC}}_{\text{TM}}$ value was obtained at the end of the TM stage in the ten runs (SD = 0). The results of this experiment are shown in Table 4. We can easily verify that the improvement obtained in DE-based registration is significant, both for the 3×3 and 6×6 OCTA images.

According to the proposed method description, the TM- and DE-based registration stages only use vessel information (FAv and OCTAv images), under the assumption that the additional use of the non-vessel information could degrade the ZNCC calculation and, therefore, erroneously bias the registration process. In order to check this hypothesis, we did a second experiment, where the TM- and DE-based registration stages are reconfigured for simultaneously using vessel and non-vessel information in the FA and OCTA images. Table 5 shows the results obtained. Compared with Table 4, we can directly verify that, in both stages and regardless of whether the OCTA image is 3×3 or 6×6, the

correlation is higher when the vessel information is uniquely considered.

Although the correlation value (ZNCC) has proved to be a useful measure to guide the registration process, we will never be sure whether a correlation value above a certain fixed threshold will correspond to a right registration unless it is checked manually. Therefore, in order to objectively measure the performance of the proposed method, we proceeded as indicated below. First, we manually selected four pairs of corresponding points for each pair of FA and 3×3 OCTA images. The four points in each OCTA image are chosen in such a way that each of them belongs to a different quadrant of the aforementioned image. Then, the four corresponding points in the FA image are chosen accordingly. We also repeated the process for each pair of FA and 6×6 OCTA images.

Once the corresponding point dataset is built, the registration quality for the i-th pair of FA and OCTA images is evaluated computing the mean Euclidean distance (MDE):

$$\text{MED}_i = \frac{\sum_{j=1}^{N_p} D(\mathbf{x}_j, \mathbf{y}_j)}{N_p} \tag{8}$$

where $D(\mathbf{x}_j, \mathbf{y}_j)$ is the Euclidean distance between the pair of corresponding points $\mathbf{x}_j = (x_{1j}, x_{2j})$ and $\mathbf{y}_j = (y_{1j}, y_{2j})$, with $\mathbf{x}_j$ belonging to the FA image and $\mathbf{y}_j$ belonging to the registered OCTA image, and $N_p = 4$ is the number of pairs of corresponding points. Here, we established a threshold $Th_{MED} = 300$ μm, in such a way that if $MED_i \leq Th_{MED}$, then the registration obtained for the i-th pair of FA and OCTA images was considered a success. Note that the threshold chosen is equivalent to 10% and 5% of the width of the square region covered by the 3×3 and 6×6 OCTA images, respectively. Applying this criterion, we obtained a success rate (SR) of 98.8% or, in other words, two errors in the 172 pairs of FA and OCTA images included in our dataset. The same two errors were obtained in each of the 10 runs. This coherent behavior is explained because the transformed OCTA image obtained in the TM-based registration stage is very far from its right registration. Therefore, the DE-based registration stage cannot do anything relevant because its radius of action is limited to a neighborhood region of the TM-based registration. In Fig. 7, these two wrong registrations are shown. Note that in both cases the wrong registration occurred with $3 \times$
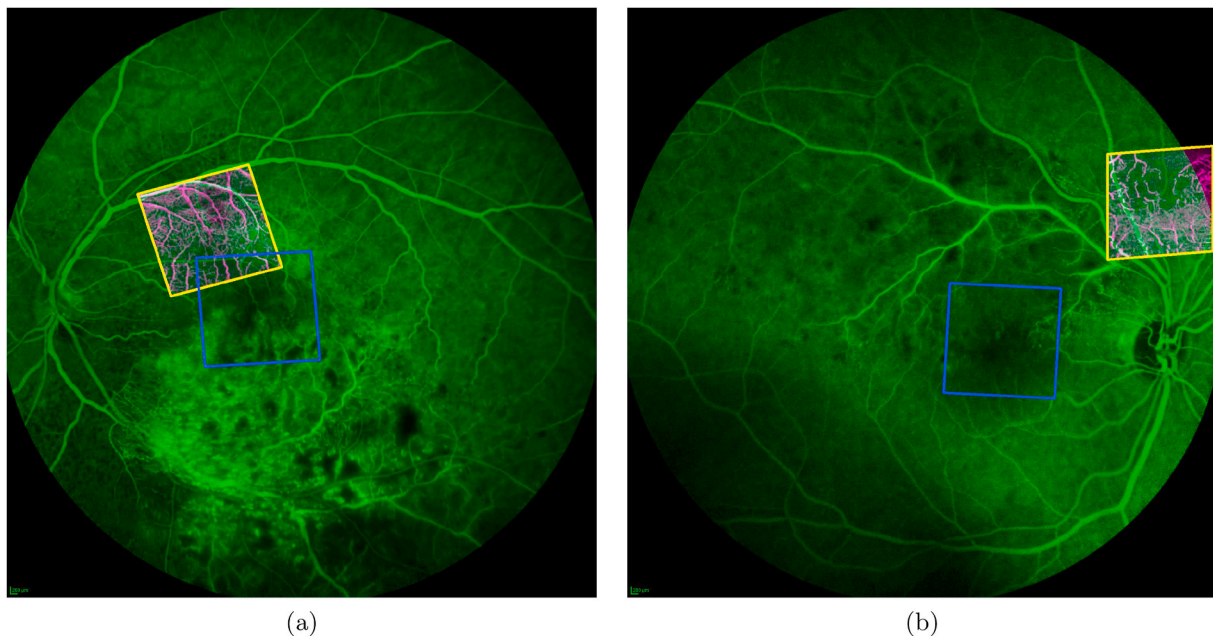
Fig. 7. The two wrong registrations obtained by our method when it was applied to the entire dataset. In both cases, the error occurred registering FA and 3×3 OCTA images: (a) the eye is pathological and OCTA image does not have artifacts; (b) the eye is pathological and OCTA image has artifacts. The right registration is shown by the blue line square.
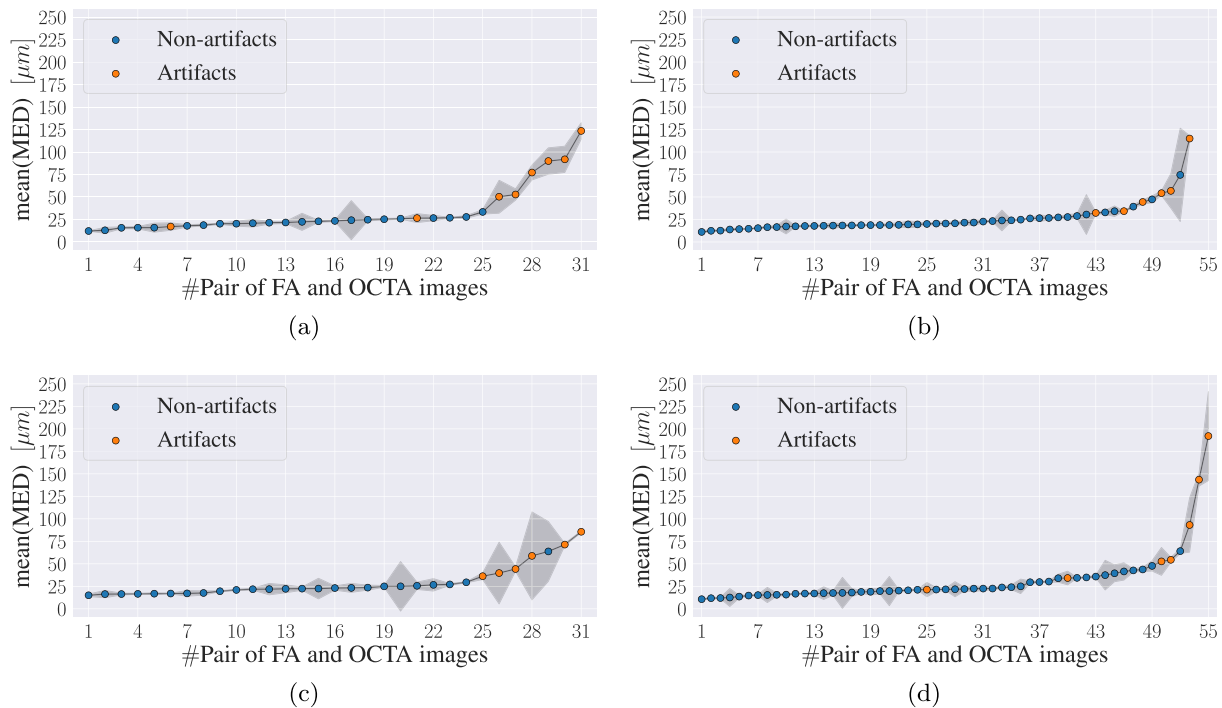
Fig. 8. Registration results for each pair of FA and OCTA images expressed in terms of the mean and standard deviation (gray areas) of the MED values obtained in the ten runs, considering different scenarios: (a) 3 × 3 OCTA in healthy eyes; (b) 3 × 3 OCTA in pathological eyes; (c) 6 × 6 OCTA in healthy eyes; and (d) 6 × 6 OCTA in pathological eyes.

3 OCTA images, in pathological eyes (in which the challenge is greater) and, additionally, the area corresponding to the right registration was characterized by the almost total absence of vessels in the FA image (here we must remember that our method only uses vessel-based information). From now on, to calculate the rest of the metrics presented in this work, the two pairs of images that showed an erroneous registration will not be considered, that is, they are considered outliers.

We can compare our method with other automatic registration methods in the related literature. However, as already mentioned in the Introduction section, the only known approach is the one described in Ref. [10]. The SR obtained there was equal to 87.5% (21 of 24 registrations work well). However, no more details related to the quality of the registration were provided by its authors.

The sequence of mean values for the MED values obtained in the ten runs in each pair of FA and OCTA images is shown and ordered from least to greatest in Fig. 8. The standard deviation associated with each mean value is also shown. We can see, as expected, that those registrations, where OCTA images with artifacts are involved, have greater error and dispersion, regardless of whether the patient's eye is healthy or pathological or whether the OCTA image is 3 × 3 or 6 × 6. Also note that, whether the OCTA image does not have any artifacts, the MED value is always below 75 μm in all the four scenarios mentioned above. In any case, it can easily be verified that the MED value obtained for any pair of images is always below the previously defined threshold Th_{MED}.

Additionally, to evaluate the behavior of the proposed method in a more compact way, we computed the average of the MED_i values obtained in the entire image dataset:

$$\overline{\text{MED}} = \frac{\sum_{i=1}^{N_T} \text{MED}_i}{N_T} \qquad (9)$$

being N_T the total number of FA and OCTA image pairs. However, given the stochastic nature of an EA, we finally calculated the mean, $\mu_{\overline{\text{MED}}}$, and SD of the $\overline{\text{MED}}$ values obtained after the ten different runs. Table 6 shows these results considering the different scenarios described in our dataset. We can observe that the method's behavior is quite consistent

Table 6
Results of averaging the $\overline{\text{MED}}$ values for the ten different runs of the proposed method. All the values are in μm. The notation $\mu \pm \sigma$ expresses the mean of the $\overline{\text{MED}}$ values and its deviation standard.

	3 × 3 OCTA		6 × 6 OCTA	
	no-artifacts	artifacts	no-artifacts	artifacts
Healthy	21.64 ± 1.11	66.15 ± 4.33	23.09 ± 1.42	56.03 ± 11.30
Pathological	22.75 ± 1.54	56.18 ± 3.86	24.40 ± 0.61	84.52 ± 10.35
All	22.39 ± 1.00	61.87 ± 2.88	23.95 ± 0.70	71.37 ± 6.87

when there are not artifacts in the OCTA images, regardless of whether these images are 3 × 3 or 6 × 6 or whether the eye is healthy or pathological. This consistency is revealed in two ways: first, the mean values are small and almost equal to the ones in the scenarios above mentioned and, second, the dispersion of the results produced by the EA is very low. On the other hand, when there is presence of artifacts in the OCTA images, the errors in the mean values almost triple. In any case, the worst case, $\mu_{\overline{\text{MED}}} = 84.52$ μm, only represents 2.8% and 1.4% of the width of the square region covered by the 3 × 3 and 6 × 6 OCTA images, respectively. Lastly, in order to visualize the behavior of the proposed method, Fig. 9 shows registration examples obtained in different pairs of FA and OCTA images, each of them illustrating different scenarios. Once again, the pairs of corresponding points can help to visually assess the quality of the registration examples that are shown. Each subfigure shows the region of interest where the registration was made. In particular, Fig. 9 (f) and (l) show the worst scenario, that is, when 3 × 3 and 6 × 6 OCTA images with artifacts are implied in the registration process, respectively. These artifacts are associated with different vessel discontinuities in the OCTA images (see Fig. 1 (e) and (f), respectively).
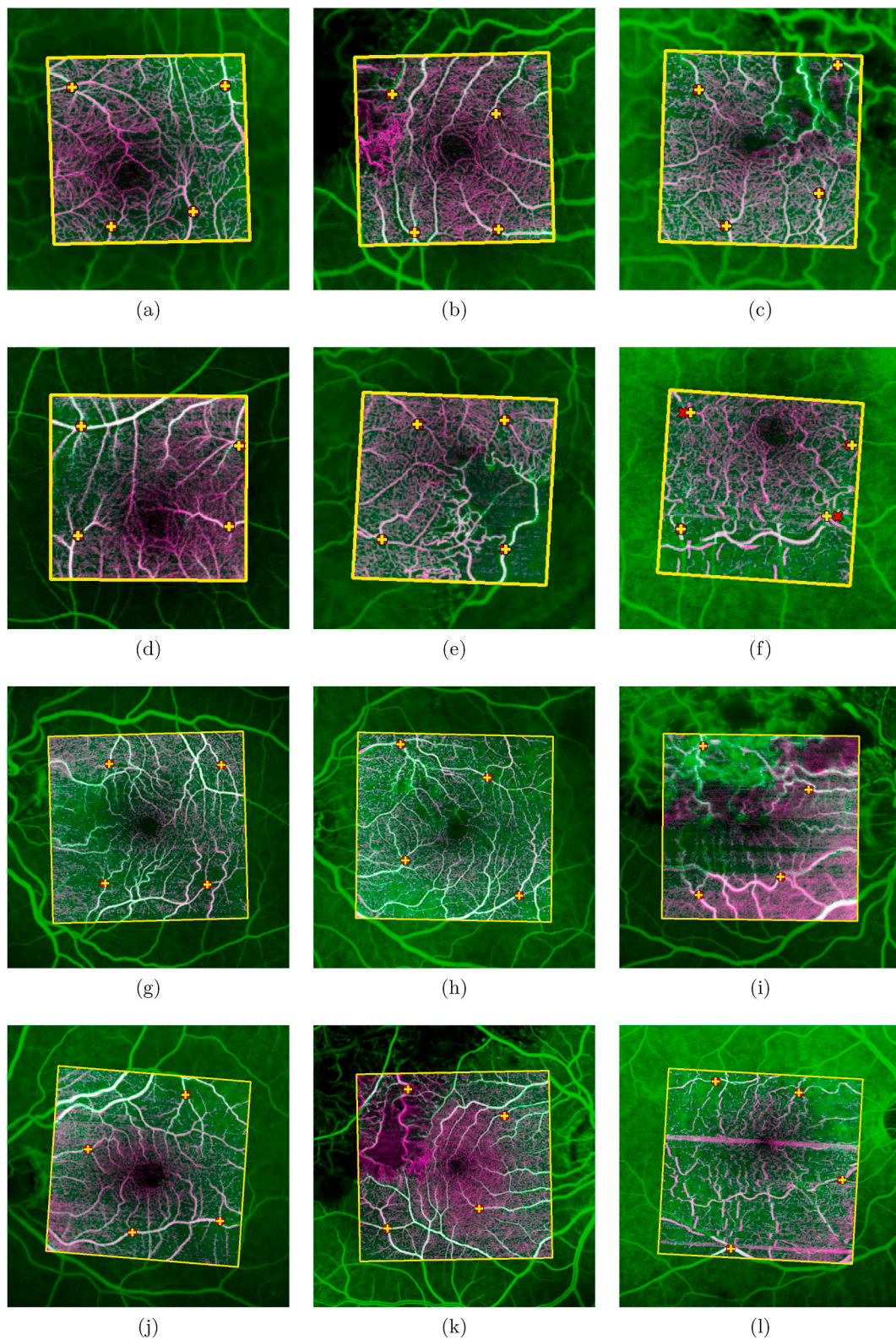
Fig. 9. Several output examples of the registration method: (a)–(f) FA and 3×3 OCTA images; (g)–(l) FA and 6×6 OCTA images. Each pair of corresponding points is represented by the symbols $+$ and $\times$.

4.2. Impact of the preprocessing parameters

To study the effect of the preprocessing parameters in the performance of our method, we proceeded in the following manner: the parameters associated with the same type of operator were varied in a neighborhood of their tuned values, while the parameters of the rest of the operators remained fixed and equal to their respective tuned values (see Table 2). Then, the effect of each combination of values was analyzed.

Thus, we considered the sets {17, 19, 21} and {3, 5, 7} pixels for the

Table 7

Results of averaging the $\overline{MED}$ values for the ten runs of a variant of the proposed method that uses half resolution in the FAv and OCTAv images when the DE-based registration stage is applied. All the values are in μm. The notation $\mu \pm \sigma$ expresses the mean of the $\overline{MED}$ values and its deviation standard.

	3 × 3 OCTA		6 × 6 OCTA	
	no-artifacts	artifacts	no-artifacts	artifacts
Healthy	23.22 ± 1.41	69.53 ± 6.31	25.21 ± 3.98	52.63 ± 4.43
Pathological	23.50 ± 1.22	57.30 ± 4.32	24.99 ± 1.89	79.71 ± 6.59
All	23.41 ± 1.15	64.29 ± 4.04	25.07 ± 1.50	67.21 ± 4.51

Table 8

Computational cost at original and half resolution. It is expressed in terms of the mean number of generations ($\overline{N}_{Gen}$) and mean time ($\overline{T}_{DE}$) used in the DE-based registration stage and the mean total time ($\overline{T}_{total}$) spent by the method in both scenarios (resolutions). The notation $x \pm \sigma$ expresses the mean and standard deviation.

|	Original resolution			Half resolution		
	$\overline{N}_{Gen}$	$\overline{T}_{DE}$ (s)	$\overline{T}_{total}$ (s)	$\overline{N}_{Gen}$	$\overline{T}_{DE}$ (s)	$\overline{T}_{total}$ (s)
3 × 3 OCTA	34.3 ± 3.8	7.4 ± 0.8	9.1 ± 0.8	36.0 ± 5.2	3.3 ± 0.4	4.8 ± 0.4
6 × 6 OCTA	29.6 ± 2.4	7.0 ± 0.6	8.9 ± 0.6	29.8 ± 2.8	2.9 ± 0.2	4.6 ± 0.2

Gaussian and median filter sizes, respectively. By testing all the possible combinations, the maximum number of erroneous registrations was equal to 4 in the worst case. Basically, the number of errors increased above 2 when the median filter size was increased from its tuned value (all the errors were obtained with 3 × 3 OCTA images). Regarding the top-hat operator, the analyzed sets were {29, 31, 33} and {13, 15, 17} pixels for the SE radius in FA and OCTA images, respectively. In this case, we did not observe practically any variation in relation to the current number of errors (a maximum of 3 registration errors were obtained in the worst case and they were always associated with 3 × 3 OCTA images). For the standard deviation in the Gaussian kernel, we considered the sets {3.0, 3.5, 4.0} and {3.5, 4.0, 4.5} for the FA and OCTA images, respectively. Here, a maximum of 5 registration errors was obtained in the worst case (one of them with 6 × 6 OCTA image). Finally, the sets {190, 200, 210} and {210, 220, 230} were selected for

Th_{max} with FA and OCTA images, respectively, and the sets of percentages {15, 20, 25}, {10, 15, 20} and {25, 30, 35} for Th_{min} with FA, 3 × 3 and 6 × 6 OCTA images, respectively. In this case, we observed that the number of errors only increased above 2 (always with 3 × 3 OCTA images) when $Th_{min_{FA}}$, $Th_{min_{3\times3}}$, and $Th_{min_{6\times6}}$ were simultaneously less than 20, 15, and 30%, respectively. This behavior is logical: as the Th_{min} decreases, the number of segmented vessels also decreases and, therefore, the registration process should get worse, given that our method only uses vessel information.

From this study, we can affirm that the search of the optimal pre-processing parameter configuration is not critical, and it is even easier when the 6 × 6 OCTA images are involved in the registration process.

4.3. Computational cost

Another interesting aspect is related to the computational cost of the proposed method. Since the DE-based registration stage uses an EA, the population-based search, which is used by these types of algorithms to find the solution, will impose a time penalization greater than that associated with the two previous stages (preprocessing and TM). In order to try to reduce the cost of evaluating the fitness per individual, we did another experiment where the half resolution of the FAv and OCTAv images was used in the DE-based registration stage. Table 7 shows the results of $\mu_{\overline{MED}}$ for the ten runs of the new variant of the proposed method, and Table 8 displays the computational cost of both variants (full and half resolution). First of all, comparing Tables 6 and 7, we can infer that, in all the scenarios considered in our image dataset, the behavior of the new variant and the original method is similar (the former produces $\mu_{\overline{MED}}$ values slightly higher than the latter). Additionally, as shown in Table 8, this similar behavior is achieved by reducing the average total time of the new variant to almost half of the total time that the original method takes. Note that the average number of generations is almost unchanged in both variants. Therefore, given that the mean time spent by the DE stage in the original method is more than two times greater than that spent by the new variant, the time taken for each generation in the former has to be longer than the latter. In any case, as expected, much of the total computation cost of the method corresponds to the DE-based registration stage. The provided execution times were obtained for a code implemented in Python and using a computer with a processor Intel Core i7-3770 K @ 3.5 GHz (8 GB RAM). Note that the total execution time could be reduced even more if a compiled language was used instead of an interpreted language, as is the case of Python.
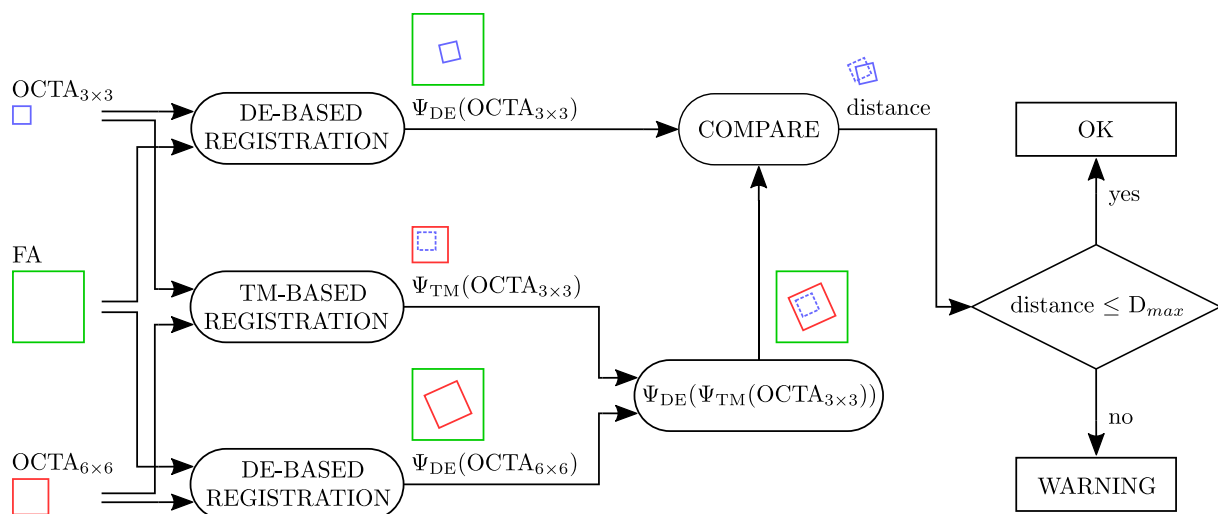
Fig. 10. Block diagram representing the procedure used to evaluate the robustness of the proposed registration method. It is assumed that, given a patient's eye, both the 3 × 3 OCTA and 6 × 6 OCTA images are available and, furthermore, the latter covers the entire area encompassed by the former. D_{max} and Ψ_X represent a distance threshold chosen by the user and a X-based affine transformation, respectively.

(a)   (b)
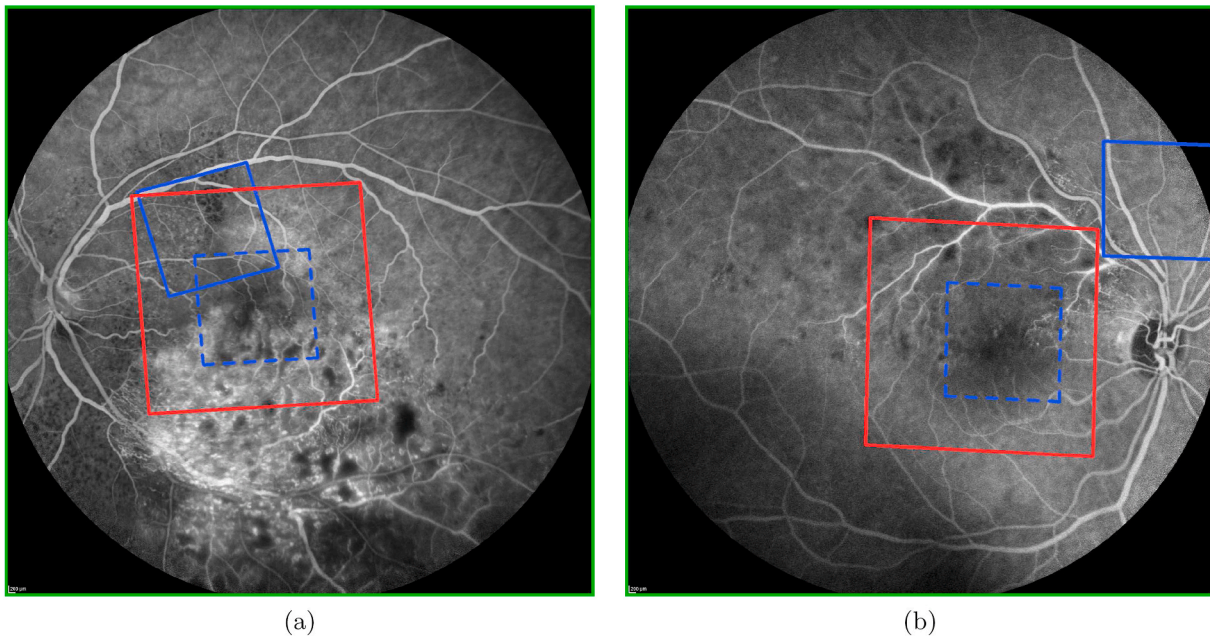
Fig. 11. The two warning cases obtained when the procedure illustrated in Fig. 10 was applied to our entire image dataset. The blue and red solid line squares represent, respectively, the registered 3×3 and 6×6 OCTA images using our registration method, and the blue dotted line square corresponds to the $\Psi_{DE}(\Psi_{TM}(OCTA_{3\times3}))$ image used as a reference for comparing.

Furthermore, in a natural way, any EA is easily parallelizable and, therefore, the DE-based registration stage could be parallelized to be executed more quickly in a computing cluster.

4.4. Automatic verification of registrations

Finally, we describe a new experiment related to the interesting advantage of working with OCT-A images of the same patient's eye at two different levels of zoom (3×3 mm and 6×6 mm). Specifically, we show that it is possible to automatically evaluate the reliability of the two registrations obtained by our method (FA-OCTA$_{3\times3}$ and FA-OCTA$_{6\times6}$). For it, we defined the following four-step procedure, which assumes that, in each pair of OCTA images, the entire area covered by the 3×3 image is always contained in the 6×6 image. In the first step, a new TM-based registration process is carried out between the 3×3 and 6×6 OCTA images. In the second step, the DE-based affine transformation, which was obtained as the solution of registering the 6×6 OCTA and FA images, is applied to the 3×3 OCTA registered image obtained in the previous step. Third, the distance between the 3×3 OCTA transformed image obtained in the second step and the 3×3 OCTA transformed image obtained using our registration method is calculated. Finally, if the distance between the centers of both images is below a threshold defined by the user, D_{max}, then it can be automatically established that the two registrations obtained by our registration method are correct. Otherwise, a warning message is shown. This procedure is summarized in Fig. 10. Note that if a warning message is obtained, there will be a problem in at least one of the three registrations (FA-OCTA$_{3\times3}$, FA-OCTA$_{6\times6}$ or OCTA$_{3\times3}$-OCTA$_{6\times6}$), but it is not possible to automatically know in which of them the error is. Otherwise, if the comparison result is "ok", then there will be a high probability that the registration method performed correctly in both registrations (FA-OCTA$_{3\times3}$ and FA-OCTA$_{6\times6}$).

Applying the described procedure to each pair of registrations obtained (FA-OCTA$_{3\times3}$ and FA-OCTA$_{6\times6}$) when our registration method

was applied to the entire dataset[1], we automatically obtained that our method performed well in all cases except two of them (see Fig. 11). These two cases coincided with the two errors that were found when we manually analyzed the results obtained by our registration method (see Fig. 7). The threshold used to evaluate each comparison was $D_{max} = 200$ μm. Note that this value represents 6.6% and 3.3% of the width of the square region covered by the 3×3 and 6×6 OCTA images, respectively.

5. Conclusions

In this work, we proposed a robust method that successfully addresses the problem of multimodal registration applied to FA and superficial plexus OCTA images. The final objective is to provide the expert clinician with the possibility of merging the complementary information from both imaging modalities. Hence, the proposed method uses TM to obtain a first approximate registration, which is later refined by a DE-based algorithm to finally find the optimal registration. Both techniques use a fitness function based on correlation to evaluate the degree of matching and, in this context, we showed that higher correlation values are obtained when the registration process uses only vessel information, instead of using all the information from the image. In 170 out of 172 pairs of FA and OCTA images, the proposed method worked correctly (SR = 98.8%). Regardless of whether the pair of registered images belongs to a healthy or pathological eye, the method obtained high quality registrations ($\mu_{\overline{MED}} \leq 24.4\mu m$), with the quality decreasing slightly ($\mu_{\overline{MED}} \leq 84.5\mu m$) in those cases in which the OCTA images presented motion artifacts (a characteristic noise of this type of image in clinical practice). In addition, a new procedure was presented to automatically detect when the proposed method produces incorrect registration. The execution of this procedure was able to identify the two incorrect registrations produced by our registration method in the entire image dataset. One of the disadvantages of the proposed registration method lies in its computational cost ($\overline{T}_{total} \leq 9.1$ s). To reduce this cost, a new variant was proposed in which the DE-based registration stage

---

[1] One of the cases was excluded from this experiment, given that its 3×3 OCTA image was not completely contained in its respective 6×6 OCTA image.

used half the image resolution that was used by the original method. With this new variant, it was possible to preserve the quality of the registrations obtained by the original method and reduce its average execution time to almost half.

The lines of future work can be oriented in two directions. On the one hand, given that the presence of the typical motion artifacts in OCTA images can difficult the registration process, a previous preprocessing stage could be applied to eliminate this kind of artifacts. However, according to the attempts made in the related literature, this stage is not immediate. On the other hand, we also think that the availability of paired and aligned OCTA and FA images presents enormous potential for exploitation by other computational techniques, such as those based on deep learning. Assuming that these techniques could produce more accurate results, they could also address more complex challenges, such as finding early signs of pathology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

[1] T. de Carlo, A. Romano, N. Waheed, J. Duker, A review of optical coherence tomography angiography (OCTA), International Journal of Retina and Vitreous 1 (2015) 1–15.

[2] D.M. Schwartz, J. Fingler, D.Y. Kim, R.J. Zawadzki, L.S. Morse, S.S. Park, S. E. Fraser, J.S. Werner, Phase-variance optical coherence tomography: a technique for noninvasive angiography, Ophthalmology 121 (1) (2014) 180–187.

[3] D. Matsunaga, J. Yi, C.A. Puliafito, A.H. Kashani, OCT angiography in healthy human subjects, Ophthalmic Surgery, Lasers and Imaging Retina 45 (6) (2014) 510–515.

[4] Y. Jia, S.T. Bailey, D. Wilson, O. Tan, M. Klein, C. Flaxel, B. Potsaid, J. Liu, C. Lu, M. Kraus, J. Fujimoto, D. Huang, Quantitative optical coherence tomography angiography of choroidal neovascularization in age-related macular degeneration, Ophthalmology 121 (7) (2014) 1435–1444.

[5] Y. Jia, S. Bailey, T. Hwang, S. McClintic, S. Gao, M. Pennesi, C. Flaxel, A. Lauer, D. Wilson, J. Hornegger, J.-G. Fujimoto, D. Huang, Quantitative optical coherence tomography angiography of vascular abnormalities in the living human eye, Proc. Natl. Acad. Sci. Unit. States Am. 112 (18) (2015) 2395–2402.

[6] M. Teussink, M. Breukink, M. van Grinsven, C. Hoyng, B. Klevering, C. Boon, E. de Jong, T. Theelen, OCT angiography compared to fluorescein and indocyanine green angiography in chronic central serous chorioretinopathy, Invest. Ophthalmol. Vis. Sci. 56 (9) (2015) 5229–5237.

[7] M. Peres, R. Kato, V. Kniggendorf, E. Cole, S. Onal, E. Torres, R. Louzada, R. Belfort, J. Duker, E. Novais, C. Regatieri, Comparison of optical coherence tomography angiography and fluorescein angiography for the identification of retinal vascular changes in eyes with diabetic macular edema, Ophthalmic Surgery, Lasers and Imaging Retina 47 (2016) 1013–1019.

[8] J. Hu, J. Qu, Z. Piao, Y. Yao, G. Sun, M. Li, M. Zhao, Optical coherence tomography angiography compared with indocyanine green angiography in central serous chorioretinopathy, Sci. Rep. 9 (2019) 1–9.

[9] M. Stattin, A.-M. Haas, A. Ahmed, U. Stolba, A. Graf, K. Krepler, S. Ansari-Shahrezaei, Detection rate of diabetic macular microaneurysms comparing dye-based angiography and optical coherence tomography angiography, Sci. Rep. 10 (2020) 1–8.

[10] R. Told, G. Reiter, A. Orsolya, T. Mittermüller, K. Eibenberger, F. Schlanitz, M. Arikan, A. Pollreisz, S. Sacu, U. Schmidt-Erfurth, Swept source optical coherence tomography angiography, fluorescein angiography, and indocyanine green angiography comparisons revisited: using a novel deep-learning-assisted approach for image registration, Retina 40 (2020) 2010–2017.

[11] S. Klein, M. Staring, K. Murphy, M.A. Viergever, J.P.W. Pluim, elastix, A toolbox for intensity-based medical image registration, IEEE Trans. Med. Imag. 29 (1) (2010) 196–205.

[12] Z. Yang, V. Vegh, D. Reutens, A fast multi-resolution differential evolution method for multimodal image registration, in: 2012 5th International Congress on Image and Signal Processing, 2012, pp. 804–809.

[13] L. Rundo, A. Tangherloni, C. Militello, M.C. Gilardi, G. Mauri, Multimodal medical image registration using particle swarm optimization: a review, in: 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 2016, pp. 1–8.

[14] Y. Chen, F. He, H. Li, D. Zhang, Y. Wu, A full migration bbo algorithm with enhanced population quality bounds for multimodal biomedical image registration, Appl. Soft Comput. 93 (2020) 106335.

[15] H. Dida, F. Charif, A. Benchabane, Grey wolf optimizer for multimodal medical image registration, in: 2020 Fourth International Conference on Intelligent Computing in Data Sciences (ICDS), 2020, pp. 1–5.

[16] P. Zang, G. Liu, M. Zhang, C. Dongye, J. Wang, A.D. Pechauer, T.S. Hwang, D. J. Wilson, D. Huang, D. Li, Y. Jia, Automated motion correction using parallel-strip registration for wide-field en face OCT angiogram, Biomed. Opt Express 7 (7) (2016) 2823–2836.

[17] R. Storn, K. Price, Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces, J. Global Optim. 11 (4) (1997) 341–359.

[18] K. Price, R.M. Storn, J.A. Lampinen, Differential Evolution: A Practical Approach to Global Optimization, Springer-Verlag, Berlin, Heidelberg, 2005.

[19] M. Díaz, J. de Moura, J. Novo, M. Ortega, Automatic wide field registration and mosaicking of OCTA images using vascularity information, Procedia Computer Science 159 (2019) 505–513, https://doi.org/10.1016/j.procs.2019.09.205.

[20] S. Sun, H.W. Park, D.R. Haynor, Y. Kim, Fast template matching using correlation-based adaptive predictive search, Int. J. Imag. Syst. Technol. 13 (2003) 169–178.

[21] X. Wang, X. Wang, L. Han, A novel parallel architecture for template matching based on zero-mean normalized cross-correlation, IEEE Access 7 (2019) 186626–186636.

[22] S. Das, P.N. Suganthan, Differential evolution: a survey of the state-of-the-art, IEEE Trans. Evol. Comput. 15 (1) (2011) 4–31.

[23] M. Saraswat, K.V. Arya, H. Sharma, Leukocyte segmentation in tissue images using differential evolution algorithm, Swarm and Evolutionary Computation 11 (2013) 46–54.

[24] P. Mesejo, O. Ibañez, O. Cordón, S. Cagnoni, A survey on image segmentation using metaheuristic-based deformable models: state of the art and critical analysis, Appl. Soft Comput. 44 (2016) 1–29.

[25] E.J. Carmona, J.M. Molina-Casado, Simultaneous segmentation of the optic disc and fovea in retinal images using evolutionary algorithms, Neural Comput. Appl. 33 (6) (2021) 1903–1921.

[26] H.-G. Beyer, The Theory of Evolution Strategies, Springer-Verlag, Berlin, Heidelberg, 2001.

[27] J.H. Holland, Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control and Artificial Intelligence, MIT Press, Cambridge, MA, USA, 1992.

[28] N. Hansen, A. Ostermeier, Completely derandomized self-adaptation in evolution strategies, Evol. Comput. 9 (2) (2001) 159–195.

[29] J. Kennedy, R.C. Eberhart, Swarm Intelligence, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2001.

[30] Z.-Z. Liu, W. Li, M. Yang, Two general extension algorithms of Latin hypercube sampling, Math. Probl Eng. (2015) 1–9.