

Correspondence

Multimodal Registration of Retinal Images Using Self Organizing Maps

George K. Matsopoulos*, Pantelis A. Asvestas,
Nikolaos A. Mouravliansky, and Konstantinos K. Delibasis

Abstract—In this paper, an automatic method for registering multimodal retinal images is presented. The method consists of three steps: the vessel centerline detection and extraction of bifurcation points only in the reference image, the automatic correspondence of bifurcation points in the two images using a novel implementation of the self organizing maps and the extraction of the parameters of the affine transform using the previously obtained correspondences. The proposed registration algorithm was tested on 24 multimodal retinal pairs and the obtained results show an advantageous performance in terms of accuracy with respect to the manual registration.

Index Terms—Multimodal image registration, ophthalmology, retina, self organizing maps.

I. INTRODUCTION

Ophthalmologists commonly compare a Red-Free (RF) retinal image, which is a reference image taken without intravenous injection of a dye, while illuminating the retina with a green light, with the corresponding Fluorescein Angiogram (FA) or the Indocyanine Green angiography (ICG) images, acquired at different times. The comparison of RF with FA or ICG retinal images is required in order to identify dynamic aspects of the circulation and evaluate a wide variety of retinal vascular disorders. The relative study of retinal images enhances the information on the reference RF images by superimposing useful information contained in FA or ICG retinal images and it is considered an important step towards a carefully directed laser treatment; a process that is commonly used in the clinical practice. The comparison of multimodal retinal images is a very difficult task due to the misalignment of the images caused by the geometry during the acquisition at different times and the possible progression of various diseases. Thus, the registration of retinal images is the key process to accurately combine information from different imaging modalities.

Automatic registration techniques have been developed to overcome failures due to application of human-interactive registration. An automatic registration algorithm of RF and FA images was presented in [1]. The algorithm was based on vessels detection, extraction of bifurcation points and the utilization of a Bayesian Hough transform for the matching. Point correspondence of segmented vessels was also used as a first step in a comparative study towards the establishment of an automatic retinal registration scheme [2]. Three transformation

models (affine, bilinear, and projective) as well as three optimization methods (downhill simplex, simulated annealing, and genetic algorithms) were evaluated in terms of accuracy and efficiency by registering RF with corresponding FA or ICG retinal images. The experimental results showed the superiority of genetic algorithms, as a global optimization process, in conjunction with the affine or bilinear transformation model. Similarly, in a recent evaluation study, three transformation types (similarity, affine and second-order polynomial) have been tested to correctly match bifurcation points selected as control points along with fourteen pixel-level fusion techniques [3]. The quantitative results on retinal images from different modalities, different resolutions and different times showed an advantageous performance of the affine transformation for the majority of the data, in terms of transformation selection and the laplacian method, in terms of fusion of the registered data. Image similarity measures, such as the mutual information, and simulating annealing with pyramid sampling, as search method, have been also used to provide robust registration under large transformations between the images and significant changes in light intensity [4].

The aim of this paper is to provide a general framework for registering multimodal retinal images (RF with corresponding FA or ICG images) by means of the theory of the self organizing maps (SOMs) network. The proposed method extends beyond what is published before in two main aspects: First, bifurcation points are extracted only on the reference image, and second, a novel implementation of the SOM network is utilized in order to determine the corresponding points on the FA or ICG image.

II. SUBJECTS AND PROCEDURES FOR ACQUISITION OF RETINA IMAGES

Retinal images were acquired using the IMAGEnet 1024 system, which is a fully functional digital imaging system for acquisition, analysis, storage, and retrieval of retinal images. Digital RF, FA, and ICG images of size 1024×1024 pixels and pixel size of about $10 \mu\text{m}$ were directly obtained using a charge-coupled device (CCD) camera that was mounted on the Topcon TRC-50IX, providing 50° angle of coverage, 39-mm working distance, and special filters for FA and ICG. The FA images were acquired in a time interval of 1–2 min after the administration of the dye (sodium fluorescein). This time interval corresponds to the mid phase of the angiogram (also known as the recirculation phase) and was chosen in order to avoid the laminar effect. During this phase, the veins and arteries remain roughly equal in brightness and thus having the same appearance in the FA image. The intensity of fluorescence diminishes slowly during this phase as much of the fluorescein is removed from the bloodstream on the first pass through the kidneys. Furthermore, control images were taken before injection of fluorescein to detect the possible presence of pseudofluorescence and/or autofluorescence [5]. Only cases without the presence of pseudofluorescence and/or autofluorescence were included in the study. The ICG images were acquired about 15 min after dye administration (indocyanine). During this time interval, retinal vessels are still visible and are not obscured by the choroid circulation.

For each case, a series of images was acquired within the aforementioned time intervals and an expert selected the images to be included in the present study. Prior to any intravenous injection, a RF image was acquired using a green filter, which causes the retinal blood vessels to appear dark.

Manuscript received February 15, 2004; revised August 4, 2004. This work was supported by the Greek Ministry of Education and Religion Affairs under the PYTHAGORAS Program. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was W. Higgins. Asterisk indicates corresponding author.

*G. K. Matsopoulos is with the Institute of Communications and Computer Systems, National Technical University of Athens, 15780 Zografos, Greece (e-mail: gmatso@esd.ece.ntua.gr).

*P. Asvestas, N. A. Mouravliansky, and K. K. Delibasis are with the Institute of Communications and Computer Systems, National Technical University of Athens, 15780 Zografos, Greece.

Digital Object Identifier 10.1109/TMI.2004.836547

The selected retinal images were driven from the CCD camera to a personal computer (Pentium 4 1.8 GHz with 512 MB RAM), where the developed automatic registration algorithm is currently running. A total of 24 cases were used in the present study; 18 cases with RF-FA images (Pair-1 to Pair-18) and 6 cases with RF-ICG images (Pair-19 to Pair-24). The cases included healthy and nonhealthy retinas (early stages of diabetic retinopathy, ischemic neuropathy, early stages of age-related macular degeneration).

III. THE PROPOSED REGISTRATION ALGORITHM

The proposed registration algorithm comprises the following steps:

- vessel centerline detection and extraction of bifurcation points in the reference image (RF image);
- automatic correspondence using a modification of the SOMs;
- extraction of the parameters of the affine transform using the previously obtained correspondences.

A. Vessel Centerline Detection and Bifurcation Point Extraction

Several methods for vessel detection and bifurcation extraction have been proposed in the literature. In [1], the extraction of the vascular tree is based on the sign of the crosscurvature and supremums of openings by means of linear structuring elements. The obtained vessels are thinned to one pixel width and the bifurcation points are detected using a supremum of openings with revolving structuring elements with a T shape. Bifurcation points that are found to be close are grouped and may become a trifurcation. In [3], the vessel centerlines are extracted by means of binarization of the retinal image using a Gaussian-constant false-alarm rate threshold and pixels in the obtained binary image with three or four neighbors are labeled as bifurcation points. In [6], the utilization of a matched filter response in conjunction with piecewise threshold probing is adopted in order to locate blood vessels in retinal images. In [7], a recursive vasculature tracing algorithm is used to map out the blood vessel structure and detect bifurcations in retinal image.

In our implementation, the vessel centerlines are detected by means of differential geometry, as follows. For a gray level retinal image $I(x, y)$, the centerlines of vessels that are brighter (darker) than the surrounding background consist of the ridge (gorge) points of the surface $S = \{(x, y, I(x, y))\}$. The detection of the ridge (gorge) points can be accomplished using the following methodology: First, the second derivatives $I_{xx}(x_0, y_0)$, $I_{xy}(x_0, y_0)$, and $I_{yy}(x_0, y_0)$ of the image with respect to x and y at each pixel position (x_0, y_0) are estimated by convolving the image with the appropriate two-dimensional (2-D) Gaussian kernels [8]. Then, the eigenvalues and eigenvectors of the Hessian matrix

$$\mathbf{H}(x_0, y_0) = \begin{bmatrix} I_{xx}(x_0, y_0) & I_{xy}(x_0, y_0) \\ I_{xy}(x_0, y_0) & I_{yy}(x_0, y_0) \end{bmatrix} \quad (1)$$

are calculated using the Jacobi method [9]. Let $\mathbf{u} = (u_x, u_y)$, $u_x^2 + u_y^2 = 1$, be the eigenvector corresponding to the eigenvalue of maximum negative (positive) value. Then, the point (x_0, y_0) is a ridge (gorge) point if the first directional derivative across the direction of \mathbf{u} at (x_0, y_0) vanishes, namely

$$\langle \nabla I(x_0, y_0), \mathbf{u} \rangle = 0 \quad (2)$$

where $\langle \rangle$ denotes the inner product. In practice, it is rather unlikely for (2) to hold exactly. One approach would be to require the absolute value of the first directional derivative across the direction of \mathbf{u} at (x_0, y_0) to be less than a predefined threshold (for example 10^{-5}). However, the accuracy of the results depends heavily on the value of the selected threshold. Therefore, the following procedure, which provides the position of a ridge (gorge) point with subpixel accuracy, is applied [8]:

First, of all, it is assumed that the pixel at position (x_0, y_0) is a ridge (gorge) point iff there are real numbers Δx and Δy such that

$$\langle \nabla I(x_0 + \Delta x, y_0 + \Delta y), \mathbf{u} \rangle = 0 \quad (3)$$

and

$$(\Delta x, \Delta y) \in [-0.5, 0.5] \times [-0.5, 0.5]. \quad (4)$$

The meaning of this assumption is that we explore a neighborhood around the current pixel in order to detect a ridge (gorge) point.

The Taylor series of the function $I(x, y)$ around the point (x_0, y_0) is given by

$$\begin{aligned} I(x, y) \approx & I(x_0, y_0) \\ & + [(x - x_0)I_x(x_0, y_0) + (y - y_0)I_y(x_0, y_0)] \\ & + \frac{1}{2} [(x - x_0)^2 I_{xx}(x_0, y_0) \\ & + 2(x - x_0)(y - y_0)I_{xy}(x_0, y_0) \\ & + (y - y_0)^2 I_{yy}(x_0, y_0)] \end{aligned} \quad (5)$$

where I_x, I_y are the first derivatives of I with respect to x and y and higher order terms of the Taylor series have been omitted. Let us assume that

$$(\Delta x, \Delta y) = (tu_x, tu_y) (t \in \mathbb{R}). \quad (6)$$

From (3) and (5), the following relation is obtained:

$$t = - \frac{u_x I_x(x_0, y_0) + u_y I_y(x_0, y_0)}{u_x^2 I_{xx}(x_0, y_0) + 2u_x u_y I_{xy}(x_0, y_0) + u_y^2 I_{yy}(x_0, y_0)}. \quad (7)$$

If the combination of the (6) and (7) satisfies (4), then the pixel at position (x_0, y_0) is declared as a ridge (gorge) point; the exact position of the ridge (gorge) point is at $(x_0 + tu_x, y_0 + tu_y)$. The magnitude of the corresponding eigenvalue quantifies the strength of the ridge (gorge) point (like the way the gradient magnitude quantifies the strength of an edge).

In order to detect the prominent vessel centerlines, a ridge (gorge) linking algorithm is applied: each ridge (gorge) pixel with ridge (gorge) strength above a threshold, T_H , is considered as a seed point of a vessel centerline. Then, a ridge (gorge) point is assigned to the centerline generated by the current seed point iff its ridge (gorge) strength is above a threshold T_L ($T_L < T_H$) and there is a path of pixels that belong to the centerline that connects this point to the seed point. The resulting vessel centerlines are thinned to one pixel width. The values of the two thresholds are image dependent. However, in general T_H lies in the range $[0.3, 0.5]$ and T_L lies in the range $[0, 0.3]$ (the values are normalized with respect to the maximum ridge (gorge) strength).

The bifurcations of the vessels are extracted using the procedure proposed in [3]. Fig. 1(a) shows the bifurcation points superimposed on a RF image. It can be seen that the bifurcation points are correctly located on the junctions of the image vessels. The vessel centerlines are shown in Fig. 1(b). The corresponding FA image, before registration, is also displayed [Fig. 1(c)]. Finally, the superposition of the vessel centerlines of the FA image on the RF image is shown in Fig. 1(d).

B. Automatic Point Correspondence Using SOMs

The SOM is a neural network algorithm, which uses a competitive learning technique to train itself in an unsupervised manner. Kohonen first established the relevant theory and explored possible applications [10]. The Kohonen model comprises of a layer of neurons m , ordered usually in a one-dimensional or 2-D grid. The training of the network is performed in an iterative way. At each iteration k , a data point $\mathbf{x} \in \mathbb{R}^n$

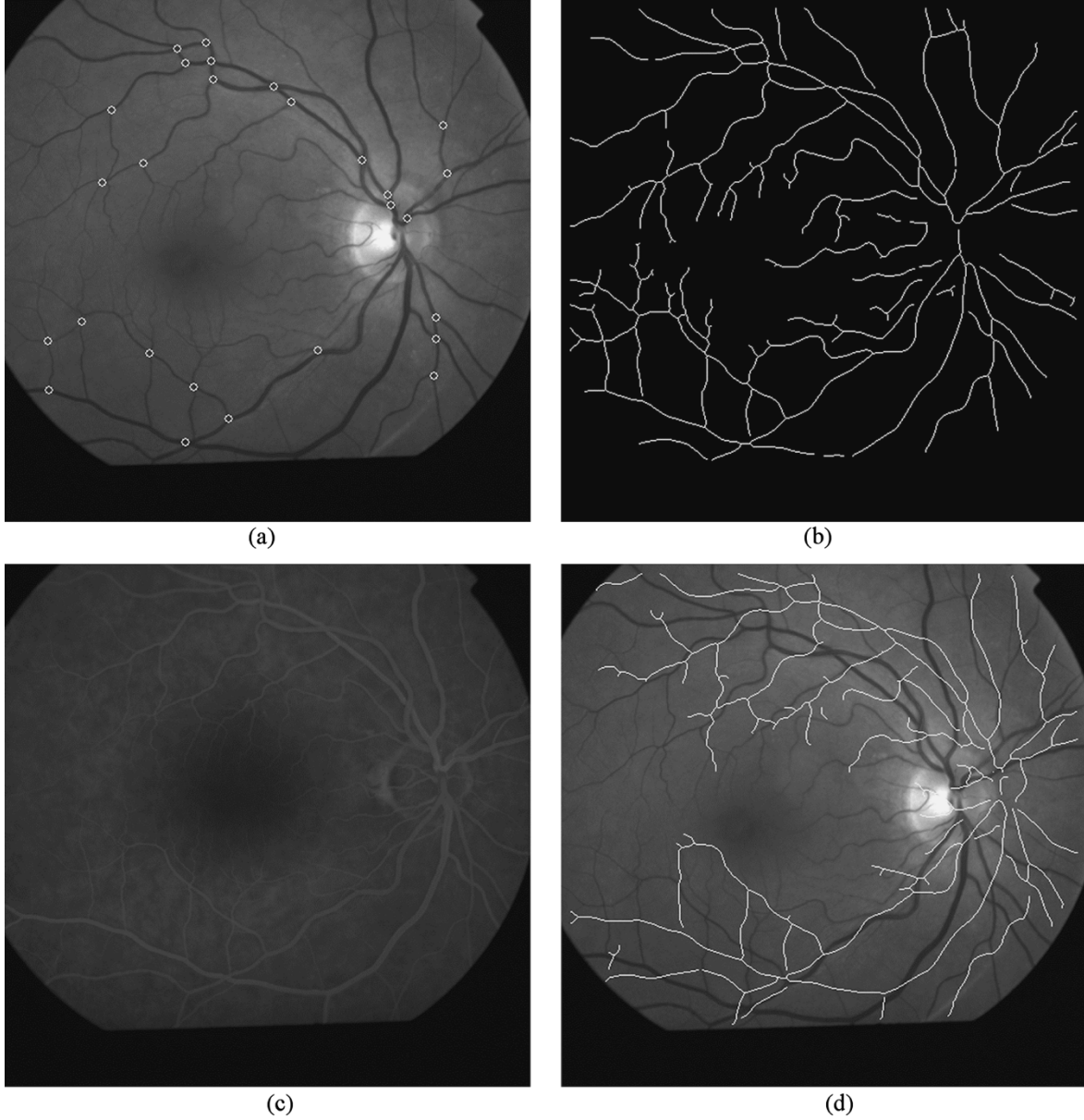


Fig. 1. (a) Extracted bifurcation points (small dots) on a RF image. (b) Vessel centerlines of the RF image. (c) The corresponding FA image (before registration). (d) Superposition of the vessel centerlines of the FA image on the RF image.

is presented to the network; the neuron j with weight vector $\mathbf{w}_j \in \mathbb{R}^n$ is declared as the winning neuron, according to the following rule:

$$j = \arg \min_i (\|\mathbf{x} - \mathbf{w}_i\|) \quad (8)$$

The winning neuron j and its neighboring neurons i have their weight vectors modified according to the following rule:

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + h_{ij}(n) [\mathbf{x}(n) - \mathbf{w}_i(n)] \quad (9)$$

where $h_{ij}(n) = h(\|\mathbf{r}_i - \mathbf{r}_j\|, n)$ is a kernel defined on the neural network space as a function of the distance $\|\mathbf{r}_i - \mathbf{r}_j\|$ between the winning neuron j and its neighboring neurons i , as well as the iteration number n . This kernel has the shape of the “Mexican hat” function, which in its discrete form has maximum value at inter-neuron distance in the case of $i = j$ whereas its value drops in a Gaussian manner as the distance increases. The width of this function decreases monotonically with iteration number. In this way convergence to the global optimum is

attempted during the early phases of the self-training process, whereas gradually the convergence becomes more local as the size of the kernel decreases.

Prior the description of the proposed method, some notations must be introduced. Let $\mu_A(I)$ denote the restriction of an image I to the region $A \subset \mathbb{R}^2$ and $T_{\mathbf{w}}(A) \subset \mathbb{R}^2$ is the rigid transformation, with parameters $\mathbf{w} = (dx, dy, \theta)$, of the region A , where dx , dy and θ are the horizontal displacement, the vertical displacement and the angle of rotation, respectively. Furthermore, $\text{MoM}(I_1, I_2)$ denotes a measure of match between two images I_1 and I_2 .

If I_R , I_F are the reference image and the image to be registered, respectively, then the implementation of the SOM network for registering the two images is as follows: The topology of the network is constructed by placing a neuron on each bifurcation point, $\mathbf{P}_i = (x_i, y_i)$ ($i = 1, 2, \dots, N$, N is the number of bifurcations), of the reference image. Each neuron is associated with a square area $A_i = [x_i - R, x_i + R] \times [y_i - R, y_i + R]$, of $(2R + 1)^2$ pixels, centered at the position of the neuron. Additionally, a weight vector $\mathbf{w}_i = (dx_i, dy_i, \theta_i)$, which

holds the parameters of a local rigid transformation, is assigned to each neuron.

The SOM network is trained as follows.

1) For each neuron, the components of the weight vector are initialized to zero values, $\mathbf{w}_i(0) = (0, 0, 0)$, the quantities $\text{MoM}_i(0) \equiv \text{MoM}(\mu_{A_i}(I_R), \mu_{T_{\mathbf{w}_i(0)}(A_i)}(I_F))$ are calculated, the variable MoM_{best} is set to a very large (in magnitude) negative value and the iteration variable, n , is set to 1.

2) While n is less than n_{max} :

- If the average value of the $\text{MoM}_i(n-1)$, $\text{MoM}_{\text{ave}}(n-1)$, is better than MoM_{best} , then $\text{MoM}_{\text{best}} = \text{MoM}_{\text{ave}}(n-1)$ and the current weights are stored as \mathbf{w}_i .
- An input signal, $\mathbf{s}(n) = (dx(n), dy(n), \theta(n))$, is generated randomly.
- For every neuron, the quantity $\text{MoM}_i(n) \equiv \text{MoM}(\mu_{A_i}(I_R), \mu_{T_{\mathbf{s}(n)}(A_i)}(I_F))$ is calculated.
- The winning neuron, k_n , in the current iteration, is defined as

$$k_n = \arg \max_i \{\text{MoM}_i(n)\} \quad (10)$$

under the condition

$$\text{MoM}_{k_n}(n) > \text{MoM}_{\text{ave}}(n-1). \quad (11)$$

- The weights of the neurons are updated according to the following equation:

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + h(k_n, n, i) [\mathbf{s}(n) - \mathbf{w}_i(n-1)] \quad (12)$$

where $h(k_n, n, i)$ ($i = 1, 2, \dots, N$) is given by the following equation:

$$h(k_n, n, i) = \begin{cases} L^{q(n)}, & \|\mathbf{P}_{k_n} - \mathbf{P}_i\| < \alpha^{q(n)} d_0 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

$$q(n) = \left\lfloor \frac{n}{p+1} \right\rfloor.$$

L , a , $d_0 \in \mathbb{R}$ and $p \in \mathbb{R}$ are parameters to be defined later, $\|\cdot\|$ denotes the Euclidean norm and $\lfloor \cdot \rfloor$ is the floor function.

- The iteration variable is increased by one.

3) When the training is finished, the parameters of the affine transformation between the two retinal images are calculated using a least squares method between the point sets $\{\mathbf{P}_i\}$ and $\{T_{\mathbf{w}_i}(\mathbf{P}_i)\}$.

Several issues regarding the proposed method should be discussed. First of all, in order to cope with the multimodality of the available images, the selected measure of match was the gradient difference, namely [11]

$$\text{MoM}(I_1, I_2) = \sum_{x,y} \frac{1}{1 + [I_{1x}(x, y) - I_{2x}(x, y)]^2} + \sum_{x,y} \frac{1}{1 + [I_{1y}(x, y) - I_{2y}(x, y)]^2} \quad (14)$$

where the subscript $x(y)$ denotes the $x(y)$ component of the gradient of the image. The rationale for selecting the aforementioned measure of match was that gradient measures have the advantage of filtering out low spatial frequency differences between the images, such as those caused by soft tissue. They also concentrate the contributions to the similarity measure on edge information, which intuitively appears sensible. The proposed measure of match employs the $1/(1+x^2)$ kernel, which makes the measure robust to thin line structures. Furthermore, a comparison of similarity measures (including cross correlation, mutual

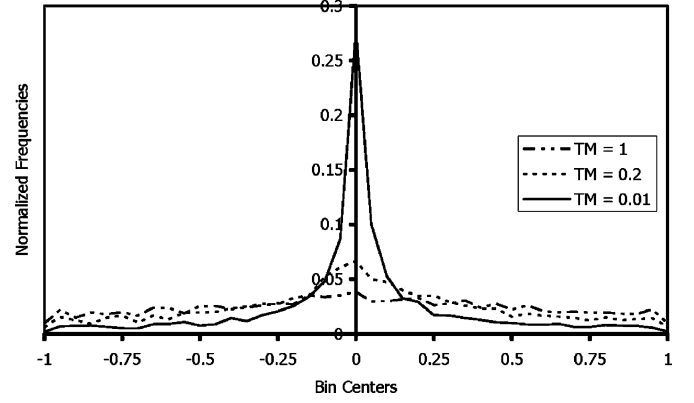


Fig. 2. Normalized histogram of the values obtained by means of the random number generator described in (15) for different values of the parameter TM .

information, gradient correlation, entropy of the difference image, gradient difference and pattern intensity) showed that the gradient difference and pattern intensity were able to register accurately and robustly even when soft-tissue structures and interventional instruments were present as differences between medical images [11].

The following generator of random numbers is used for producing the input signals to the network:

$$s_j(n) = w_{k_n, j} + \text{sgn}(v_j - 0.5) TM(n) \times \left[\left(1 + \frac{1}{TM(n)} \right)^{|2v_j - 1|} - 1 \right] \times (U_j - L_j) \quad (j = 1, 2, 3)$$

$$TM(n) = \begin{cases} 1, & n = 0 \\ \exp \left(-2 \left(q(n) \right)^{\frac{1}{p}} \right), & n > 0 \end{cases} \quad (15)$$

where $s_1(n) = dx(n)$, $s_2(n) = dy(n)$, $s_3(n) = \theta(n)$, v_j is a uniformly distributed random variable in $[0, 1]$, and U_j (L_j) denotes the maximum (minimum) allowed value for the corresponding component of the input signal. Although U_j and L_j are inputs to the registration process, for all pairs of images used in the current study, constant values were used (± 150 pixels for the displacement and $\pm 10^\circ$ for the angle of rotation).

It must be noted that (15) is a slightly modified version of the generator used in the very fast simulated annealing method [12] and provides random signals which in general lie in the range $[w_{k_n, j} - (U_j - L_j), w_{k_n, j} + (U_j - L_j)]$. When a generated signal is not in the allowed range $[L_j, U_j]$, then it is discarded and a new signal is produced until $s_j(n) \in [L_j, U_j]$. The parameter $TM(n)$ controls how far from the weights of the current winning neuron the input signal can reach. As the iteration variable evolves, the magnitude of $TM(n)$ falls exponentially and the generated input signals are more localized around the weights of the current winning neuron (see Fig. 2). This is a desired property, since as the number of iterations grows, the weights of the current winning neuron get closer to the parameters of the solution of the registration problem.

The parameter d_0 provides the initial radius of a circular region around the winning neuron. Only neurons inside this region are updated. Usually, d_0 is set to the maximum distance between bifurcation points. As can be seen from (13), this distance is reduced with geometric rate determined by the parameter α ($0 < \alpha \leq 1$). A typical value for the parameter α is 0.995. The parameter L acts like a gain constant for the magnitude of the update that is applied to the weights of the neurons. This parameter also decreases geometrically as the iteration variable evolves. The range of values L is between 0.99 and 1.0; a typical value is 0.995. The parameter p is an integer that determines the

TABLE I
ROOT MEAN SQUARE ERROR (RMSE) FOR ALL IMAGE PAIRS FOR
DIFFERENT RESOLUTIONS

Retinal Image Pairs	RMSE (in pixels)	
	512 × 512	1024 × 1024
Pair-1	1.583	2.958
Pair-2	1.182	2.160
Pair-3	1.395	2.523
Pair-4	1.296	2.365
Pair-5	1.789	3.258
Pair-6	1.145	1.989
Pair-7	1.695	3.015
Pair-8	1.569	2.945
Pair-9	0.948	1.564
Pair-10	1.689	3.025
Pair-11	1.996	3.645
Pair-12	1.863	3.487
Pair-13	1.689	3.069
Pair-14	2.012	3.956
Pair-15	2.159	4.035
Pair-16	2.023	3.845
Pair-17	1.591	3.049
Pair-18	2.109	3.963
Pair-19	1.199	2.059
Pair-20	1.839	3.523
Pair-21	1.669	3.145
Pair-22	2.389	4.245
Pair-23	1.998	3.756
Pair-24	2.301	4.256

rate of change of the parameters L and α . Practically, this parameter determines the number of iterations that are executed before an adjustment of the values for the parameters L , a , and $TM(n)$ takes place. A typical value for this parameter is 200. The number of iterations is set to 10 000 and the size of the square area associated with each neuron is 19 ($R = 9$).

It should be pointed out that a sufficient number of bifurcation points should be extracted in order to achieve an accurate registration result. Furthermore, the bifurcation points should be distributed over the whole image (if possible). The degree of sparseness of the bifurcation points can be determined by checking if the standard deviation of the $x - y$ coordinates is above a predefined threshold. Experiments have shown that six bifurcation points, with standard deviation of the $x - y$ coordinates that exceeds 100, are sufficient in order to obtain accurate registration results.

Finally, since the transformed region $T_{s(n)}(A_i)$ does not have integer coordinates, bilinear interpolation is used in order to calculate $MoM_i(n)$.

IV. RESULTS

Several methods have been proposed in the literature for the assessment of the performance of retinal image registration algorithms. The simplest approach is by means of visual inspection of a fused image. The fused image is either a checkboard of reference and transformed images or a superposition of the transformed extracted vessels on the reference image [1]. In [3], the superposition percentage was

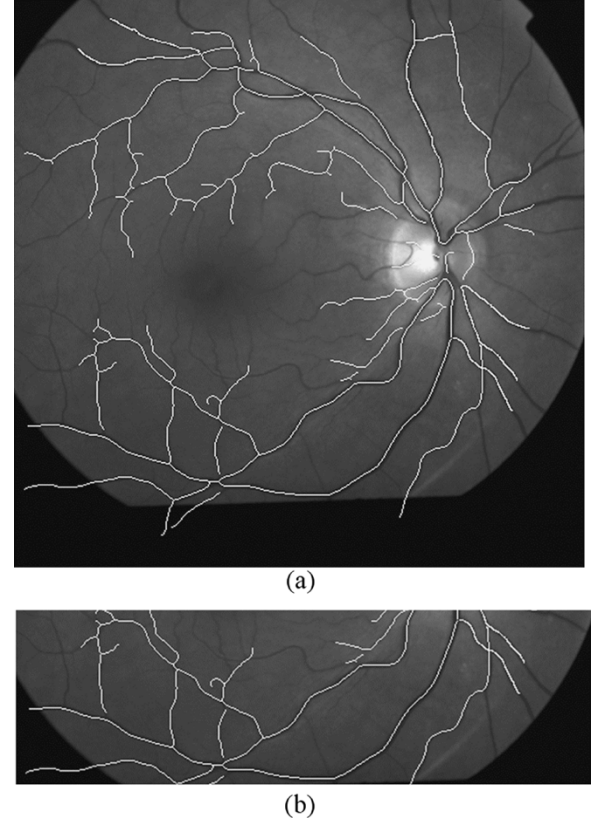


Fig. 3. (a) Superposition of the vessel centerlines of the transformed FA image on the RF image. (b) Zoomed area of the image shown in (a).

used to quantify the performance of the registration algorithm. First, the transformed extracted vessels were superimposed on the reference vessel network. The superposition percentage was calculated as the percentage of the transformed pixels that fell into a 3×3 window centered on a reference vessel network pixel. In [4], Ritter *et al.* performed an exhaustive search in the neighborhood of the results obtained through their method. The parameters of the exhaustive served as the ground truth. Then, for each transformation parameter, the root-mean-square error (RMSE), over all the available image pairs, was considered as an indicator of the accuracy of the registration.

The accuracy of the proposed registration method was quantitatively assessed for every retinal pair. Specifically, pairs of corresponding points were defined manually in both images by an experienced ophthalmologist and the parameters of the corresponding affine transformation were calculated by means of the least squares method. The error in point placement was assessed as the RMSE between the affine-transformed points and the points from the reference image. For each pair of images, the aforementioned procedure was repeated three times. The pairs of points from the trial that gave the minimum error were used for validating the proposed registration scheme. The average error in point placement was found to be within one pixel.

The validation of the proposed registration scheme was performed as follows: The points defined in the FA or ICG image were transformed using the transformation that was obtained by the proposed method and the RMSE between the transformed points and the points from the reference image was calculated. In order to ensure the consistency, the results were averaged over five independent executions of the algorithm for all retinal pairs. In order to study the behavior of the proposed method for different resolutions, the original images, with size 1024×1024 pixels, were downsampled to size 512×512 pixels using an averaging process. The results are listed in Table I. As can be seen, the RMSE was below 4 and 2 pixels (approximately $40 \mu\text{m}$) for the 1024

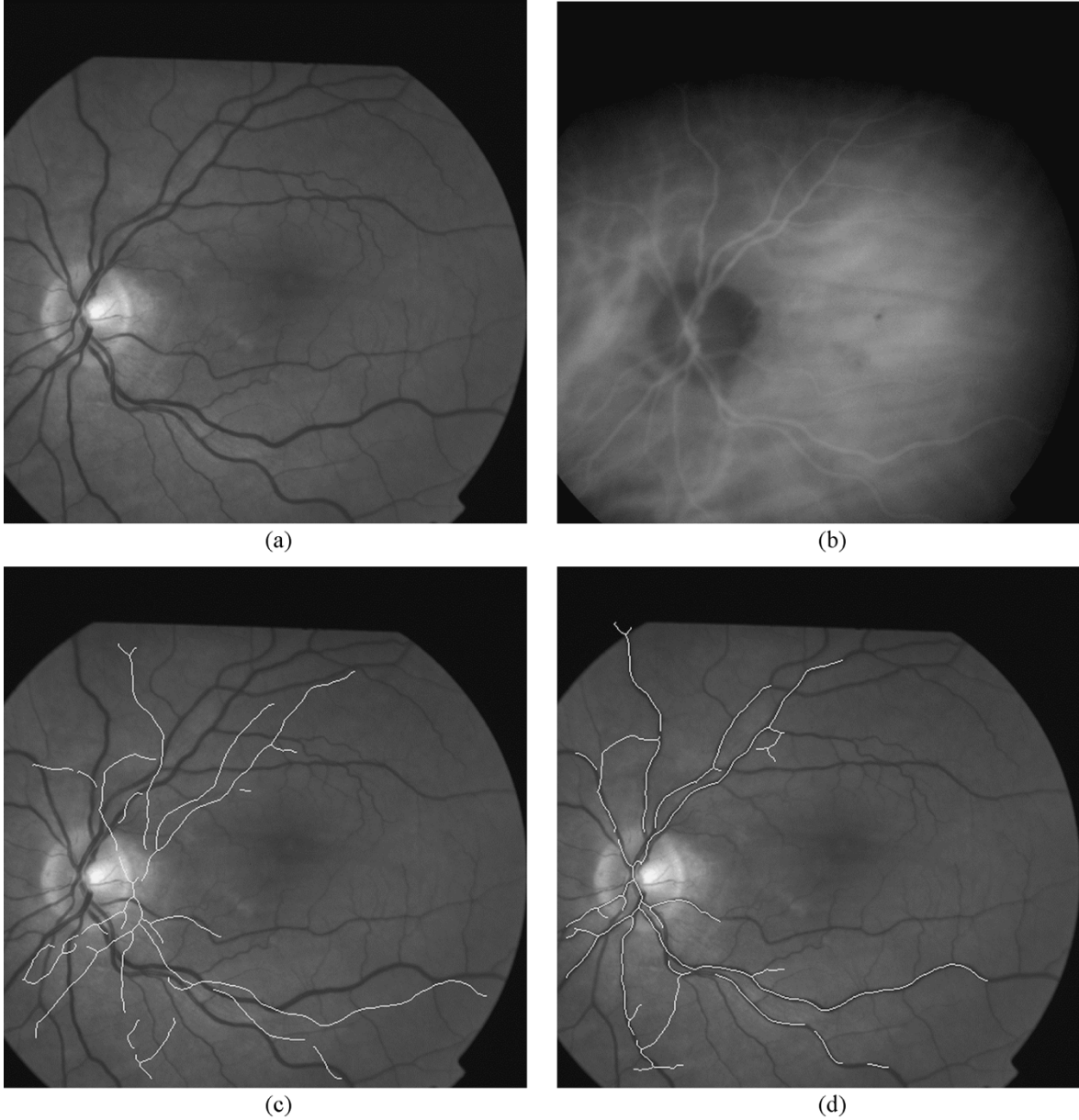


Fig. 4. (a) RF image. (b) Corresponding ICG image before registration. (c) Superposition of the vessel centerlines of the ICG image on the RF image before registration. (d) Superposition of the vessel centerlines of the ICG image on the RF image after registration.

$\times 1024$ and 512×512 images, respectively. In clinical practice, especially for laser treatment, a registration error below $50 \mu\text{m}$ is acceptable. As a rule, laser burns are created around $500 \mu\text{m}$ in diameter in the periphery and smaller in the macular region, for example $200 \mu\text{m}$ [13]. The smallest sizes 50 to $100 \mu\text{m}$ may be used for treating sub-retinal vessels near the fovea [13]. Furthermore, it can be noticed that the performance of the algorithm in terms of accuracy was superior for registering RF-FA images than RF-ICG images, due to the fact that the retinal visibility in the ICG images is inherently inferior compared to RF and FA images.

The performance of the proposed registration algorithm for the RF and FA images of Fig. 1 is visually demonstrated in Fig. 3. In particular, Fig. 3(a) shows the superposition of the vessel centerlines of the transformed FA image on the RF image. From this figure as well as from a zoomed area [Fig. 3(b)], it can be seen that a successful registration result has been achieved. Fig. 4(a) and (b) shows a pair of RF and ICG image before registration. Fig. 4(c) and (d) shows the superposition of

the vessel centerlines of the ICG image on the RF image before and after registration, respectively.

V. DISCUSSION

The following remarks can be drawn for the proposed registration scheme: It is evident that the objective function that is optimized is the average value of the partial measure of matches. A refinement of the solution is obtained at each iteration. In particular, an initial value for the parameter d_0 that is large enough is selected (usually larger or equal to the maximum distance between the available bifurcations) such that the weight vectors of all neurons are updated. In this way, the SOM network provides a rough approximation of the solution of the registration problem. As the iterations evolve, the size of the radius around the winning neuron decreases, the SOM network is self-adjusted to capture any local deformations (if present). This is accomplished in conjunction with a decrease of the value of the parameter $TM(n)$ in (15) which

controls how far from the winning neuron the input signal can reach. In order to find the global optimum of the objective function, the decrease of the size of the radius of the circular region around the position of the winning neuron should be done gradually. We have found out experimentally that a choice of p in the range 150 to 250 and α in the range 0.99 to 1.0 gives satisfactory results. Regarding the value of the parameter L , experiments have shown that it should be near unity and in particular in the range 0.995 to 1.0. It should be emphasized that the algorithm is not critically influenced by the specific chosen values of the aforementioned parameters.

The condition given by (11) is necessary to protect the SOM network from generating a signal that is much more different than the weighting vectors of the neurons in the current iteration. The condition is not too strict in the sense that it is not prohibitive of the production of a winner with a measure of match lower than that in the previous iteration. Consequently, it is achieved the escape from local optima of the objective function, since solutions that do not lead to better values of the objective function are accepted too. This approach is similar to the Metropolis process that takes place in simulated annealing [12].

Another aspect of the proposed registration scheme that should be addressed is the utilization of the gradient difference as a measure of match. The selected measure of match provided accurate results for almost all image pairs that were used in the current study. However, it should be stressed that the gradient difference may fail for images characterized by apparent hyperfluorescence. Hyperfluorescence is caused usually by a breakdown or lack of tight vascular junctions in abnormal blood vessels. Any leakage of fluorescein from a retinal vessel or within the retinal tissues indicates an abnormality. For example, capillary microaneurysms, retinal telangiectasias, arterial macroaneurysm, papilledema, and some vascularized tumors exhibit leakage.

Finally, it should be noted that the focus of the paper was to show the feasibility of the SOM theory for registering multimodal retinal images. The proposed implementation of the SOM model could be considered as a method for finding the optimum value of an objective function. Under this framework, the proposed registration scheme provides several "degrees of freedom" regarding its parameters. For example, another measure of match (such as the mutual information) could be used, the bifurcation points could be extracted by means of other methods, another transformation (such as bilinear or elastic) could be adopted after the training of the SOM network.

VI. CONCLUSION

In this paper, a new registration algorithm for registering multimodal retinal images was presented. Two basic novel implementations were introduced. First, the application of the vessel centerline detection and bifurcations extraction process only on the reference image. This step simplifies the registration methodology since candidate points are identified only on the reference image. Second, the novel implementation of the SOM network to define automatic correspondence of the bifurcation points between the reference and the image to be transformed. The proposed algorithm was tested for 24 pairs of multimodal images providing an accuracy of approximately 40 μm for all retinal pairs.

REFERENCES

- [1] F. Zana and J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," *IEEE Trans. Med. Imag.*, vol. 18, pp. 419–428, May 1999.

- [2] G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, "Automatic registration of retinal images with global optimization techniques," *IEEE Trans. Inform. Tech. Biomed.*, vol. 3, pp. 47–60, Mar. 1999.
- [3] F. Laliberte, L. Gagnon, and Y. Sheng, "Registration and fusion of retinal images: an evaluation study," *IEEE Trans. Med. Imag.*, vol. 22, pp. 404–418, May 2003.
- [4] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. van Saarloos, "Registration of stereo and temporal images of the retina," *IEEE Trans. Med. Imag.*, vol. 18, pp. 404–418, May 1999.
- [5] C. Barry, J. Singh, and I. J. Constable, "Are optic disc drusen exhibiting autofluorescence, pseudofluorescence or reflectance?," *J. Ophthalmic Photography*, vol. 22, pp. 32–35, 2000.
- [6] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, pp. 203–210, Mar. 2000.
- [7] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inform. Tech. Biomed.*, vol. 3, pp. 125–138, June 1999.
- [8] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, pp. 113–125, Feb. 1998.
- [9] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [10] T. Kohonen, *Self-Organizing Maps*. Berlin, Germany: Springer-Verlag, 2000.
- [11] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, pp. 586–595, Aug. 1998.
- [12] L. Ingber and B. Rosen, "Genetic algorithms and very fast simulated reannealing: a comparison," *Mathematical and Computer Modeling*, vol. 16, no. 11, pp. 87–100, 1992.
- [13] J. Voke, "Lasers and their use in ophthalmology: part 2," *Optometry Today—Optics Today*, pp. 31–36, June 2001.

Correction to "Pulmonary Airways: 3-D Reconstruction From Multislice CT and Clinical Investigation"

Catalin I. Fetita*, Françoise Prêteux, Catherine Beigelman-Aubry, and Philippe Grenier

In [1], (10) should have read as follows:

$$\forall x \in \text{supp}(f) \subset \mathbb{R}^n, \quad \mathcal{RC}_{f,O}^i(x, Y) \\ = \max \left(f(x), \min_{\nu \in \mathcal{V}(x)} \left(\mathcal{RC}_{f,O}^i(\nu, Y) \right) \right) \quad (10)$$

REFERENCES

- [1] C. I. Fetita, F. Prêteux, C. Beigelman-Aubry, and P. Philippe Grenier, "Pulmonary Airways: 3-D Reconstruction From Multislice CT and Clinical Investigation," *IEEE Trans. Med. Imag.*, vol. 23, pp. 1353–1364, Nov. 2004.

Manuscript received November 2, 2004. Asterisk indicates corresponding author.

*C. I. Fetita is with the ARTEMIS Project Unit, Institut National des Télécommunications, 9 rue Charles Fourier, 91011 Evry Cedex, France (e-mail: catalin.fetita@int-evry.fr).

F. Prêteux is with the ARTEMIS Project Unit, Institut National des Télécommunications, 91011 Evry Cedex, France.

C. Beigelman-Aubry and P. Grenier are with the Central Radiology Service, Pitié Salpêtrière Hospital, 75651 Paris Cedex 13, France.

Digital Object Identifier 10.1109/TMI.2004.840453