Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Registration of Retinal Angiograms Using Self Organizing Maps

P. A. Asvestas, G. K. Matsopoulos, K. K. Delibasis and N. A. Mouravliansky

Institute of Communication and Computer Systems, 9 Iroon Polytechneiou str., 15780, Athens, Greece,

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 Organized Maps (SOMs) 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- **Retinal registration, vessel centerline, automatic correspondence, Self Organized 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 Fluoroscein Angiogram (FA) images, acquired at different times. The comparison of RF with FA retinal images is required in order to identify dynamic aspects of the circulation and evaluate a wide variety of retinal vascular disorders. 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 FA and RF images was presented in [1] and [2]. In a recent evaluation study, three transformation types have been tested to correctly match bifurcation points selected as control points along with fourteen pixel-level fusion techniques [3]. 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 by means of the theory of the Self Organizing Maps (SOM) network. The proposed method extends beyond what is published before in two main aspects: Firstly, bifurcation points are extracted only on the reference image, and secondly, a novel implementation of the SOM network is utilized in order to determine the corresponding points on the FA 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 and FA images of size 1024 × 1024 pixels and pixel size of about 10μm were directly obtained using a CCD camera that was mounted on the Topcon TRC-50IX, providing 50º angle of coverage, 39 mm working distance and special filters for FA. The FA images were acquired in a time interval of 1 to 2 minutes 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). Prior to any intravenous injection, a RF image was acquired using a green filter, which causes the retinal blood vessels to appear dark. This image is the one observed by the expert during laser treatment and it is used as a reference image during the registration procedure.

The selected retinal images were driven from the CCD camera to a personal computer (Pentium 4 1.8GHz with 512MB RAM), where the developed automatic registration algorithm is currently running. A total of 18 cases were used in the present study. The cases included healthy and non-healthy 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:

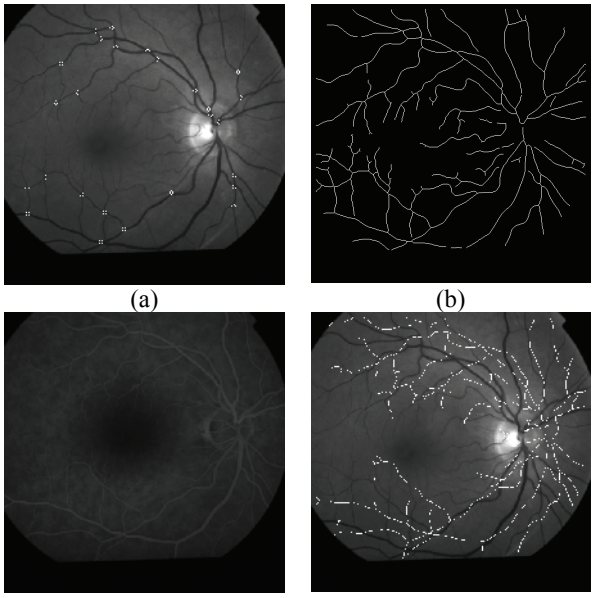
- vessel centerline detection and extraction of bifurcation points in the reference image (RF image),
- automatic correspondence using a modification of the Self Organized Maps (SOMs),
- extraction of the parameters of the affine transform using the previously obtained correspondences.

A) Vessel Centerline Detection and Bifurcation Point Extraction

Several methods for vessel detection and bifurcation extraction have been proposed in the literature. In our implementation, the vessel centerlines are detected by means of differential geometry, as described in [5].

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 if 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).



(a) (b) (c) (d)

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.

B) Automatic Point Correspondence Using SOMs

The Self-Organizing Map (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 [6].

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 $A \subset Z^2$ and $T_{\mathbf{w}}(A) \subset \Re^2$ 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, $MoM(I_1, I_2)$ denotes a measure of match between two images I_1 and I_2.

If I_R and 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, $\mathbf{P}_i = (x_i, y_i)$ ($i = 1, 2, \ldots, N$, 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 $MoM_i(0) \equiv MoM\left(\mu_{A_i}(I_R), \mu_{T_{\mathbf{w}_i(0)}(A_i)}(I_F)\right)$ 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 $MoM_i(n-1)$, $MoM_{ave}(n-1)$, is better than MoM_{best}, then $MoM_{best} = MoM_{ave}(n-1)$ 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 $MoM_i(n) \equiv MoM\left(\mu_{A_i}(I_R), \mu_{T_{\mathbf{s}(n)}(A_i)}(I_F)\right)$ is calculated.
   - The *winning neuron*, k_n, in the current iteration, is defined as:

$$k_n = \arg\max_i \{MoM_i(n)\} \quad (1)$$

   under the condition

$$MoM_{k_n}(n) > MoM_{ave}(n-1) \quad (2)$$

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

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + h(k_n, n, i)\left[\mathbf{s}(n) - \mathbf{w}_i(n-1)\right] \quad (3)$$

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

$$h(k_n, n, i) = \begin{cases} L^{q(n)}, & \left\|\mathbf{P}_{k_n} - \mathbf{P}_i\right\| < \alpha^{q(n)} d_0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

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

$L, a, d_0 \in \Re$ and $p \in Z$ are parameters to be defined later, $\|\ \|$ denotes the Euclidean norm and $\lfloor\ \rfloor$ 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 [7]:

$$MoM(I_1, I_2) = \sum_{x,y} \frac{1}{1 + \left[I_{1x}(x,y) - I_{2x}(x,y)\right]^2} + \sum_{x,y} \frac{1}{1 + \left[I_{1y}(x,y) - I_{2y}(x,y)\right]^2} \quad (5)$$

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 $\frac{1}{1+x^2}$ kernel, which makes the measure robust to thin line structures. Furthermore, a comparison of similarity measures (including cross correlation, mutual 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 [7].

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 (4), this distance is reduced with geometric rate determined by the parameter α ($0 < \alpha \le 1$). A typical value for the parameter $a = 0.995$. 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 rate of change of the parameters L and a. Practically, this parameter determines the number of iterations that are executed before an adjustment of the values for the parameters L and a 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 $T_{\mathbf{s}(n)}(A_i)$ does not have integer coordinates, bilinear interpolation is used in order to calculate $MoM_i(n)$.

IV. Results

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 root mean square error (RMSE) between the affine-transformed points and the points from the reference image. For each pair of images, the aforementioned procedure was repeated 3 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 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 5 independent executions of the algorithm for all retinal pairs and for resolutions of 1024×1024 and 512×512 pixels.

The results are listed in Table I. As can be seen, the RMSE was below 4 and 2 pixels (approximately 40μm) for

the 1024×1024 and 512×512 images, respectively. In clinical practice, especially for laser treatment, a registration error below 50μm is acceptable. As a rule, laser burns are created around 500μm in diameter in the periphery and smaller in the macular region, for example 200μm [8]. The smallest sizes 50 to 100μm may be used for treating subretinal vessels near the fovea [8].

TABLE I.

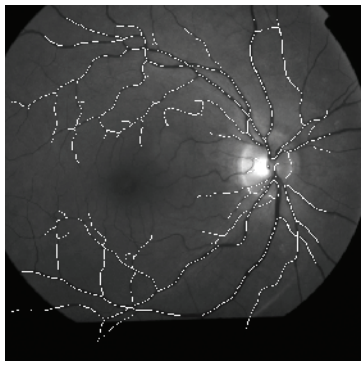
ROOT MEAN SQUARE ERROR (RMSE) FOR ALL IMAGE PAIRS FOR DIFFERENT RESOLUTIONS

	RMSE (in pixels)	
Retinal Image Pairs	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

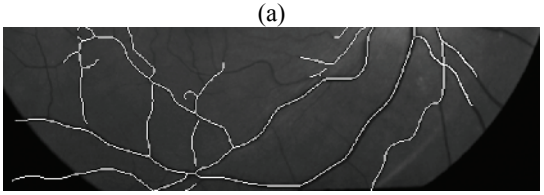
The performance of the proposed registration algorithm is also visually demonstrated in Fig. 2. In particular, Fig. 2(a) shows the superposition of the vessel centerlines of the transformed FA image on the RF image. From this figure as well as from a zoomed area (Fig. 2(b)), it can be seen that a successful registration result has been achieved.

V. CONCLUSIONS

In this paper, a new registration algorithm for registering multimodal retinal images was presented. Two basic novel implementations were introduced. Firstly, 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. Secondly, 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 18 pairs of multimodal images providing an accuracy of approximately 40μm for all retinal pairs.



(a)



(b)

Fig. 2. (a) Superposition of the vessel centerlines of the transformed FA image on the RF image. (b) Zoomed area of the image shown in (a).

ACKNOWLEDGEMENT

This work was co-funded by the European Social Fund (75%) and National Resources (25%) - Operational Program for Educational and Vocational Training II (EPEAEK II) and particularly the Program PYTHAGORAS.

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, no. 5, 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. Bio-Med.*, vol. 3, no 1, pp. 47-60, March 1999.

[3] F. Laliberte, L. Gagnon, and Y. Sheng, "Registration and fusion of retinal images: An evaluation study," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, 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, no. 5, pp. 404-418, May 1999.

[5] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. PAMI*, vol. 20, no. 2, pp. 113-125, February 1998.

[6] T. Kohonen, Self-organizing maps, Springer Verlag, Berlin, 2000.

[7] 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, no. 4, pp. 586-595, August 1998.

[8] J. Voke, "Lasers and their use in ophthalmology: Part 2," Optometry Today – Optics Today, pp. 31-36, June 2001.