

A Multimodal Registration Algorithm of Eye Fundus Images Using Vessels Detection and Hough Transform

F. Zana* and J. C. Klein

Abstract—Image registration is a real challenge because physicians handle many images. Temporal registration is necessary in order to follow the various steps of a disease, whereas multimodal registration allows us to improve the identification of some lesions or to compare pieces of information gathered from different sources. This paper presents an algorithm for temporal and/or multimodal registration of retinal images based on point correspondence. As an example, the algorithm has been applied to the registration of fluorescein images (obtained after a fluorescein dye injection) with green images (green filter of a color image). The vascular tree is first detected in each type of images and bifurcation points are labeled with surrounding vessel orientations. An angle-based invariant is then computed in order to give a probability for two points to match. Then a Bayesian Hough transform is used to sort the transformations with their respective likelihoods. A precise affine estimate is finally computed for most likely transformations. The best transformation is chosen for registration.

Index Terms—Hough transform, image registration, mathematical morphology, ophthalmology, retina.

I. INTRODUCTION

REGISTRATION of medical images is fundamental in order to follow the evolution of some lesions over the years or to compare images obtained under different conditions. This is particularly true for retinal images of eye fundus from patients developing diabetes. Different steps of most retinal diseases can be characterized by various damages of blood vessels. Temporal registration of images is necessary in order to follow the evolution of the illness, but multimodal registration is also a great challenge in the scope of an automatic diagnosis. For example, in the case of diabetic retinopathy, some articles report detection of various lesions (exudate, microaneurysms, and hemorrhages) in images taken with a green filter [1], while other articles report detection of other elements of diagnosis on angiographic images ([2]–[5]). Some vascular lesions such as microaneurysms are quite numerous [4], and counting them is so time consuming that it is impossible to track them without the help of an automatic program. Moreover, the turnover rate of microaneurysms is an essential element for foreseeing further evolution [2]. Some

different lesions may also have similar aspects, therefore, physicians often use more than one image to identify a lesion and its seriousness [3], or base their diagnosis on various detections of images of different modalities. In retina analysis, two types of images are often used for the diagnosis of the gravity of diabetic retinopathy: angiographic images obtained under ultra-violet light after the injection of fluorescein (fluorescein images) and images taken under natural light with a green filter (green images). Multimodal registration of vascular lesions is thus a fundamental task in order to identify properly each abnormality. It is complementary to the registration of images obtained from different examinations, but both are equally necessary to detect and measure the evolution of those lesions. We present an algorithm that is designed for multimodal as well as temporal registration. Based on a similar detection of the vascular tree, this algorithm could easily become part of a retinal analysis program.

A. Some Methods of Registration

Registration is a difficult task in image processing because the correspondence problem is not straightforward. As a consequence, it is not surprising that the algorithms for registration are very numerous, and each generally applies to a specific type of images. They are used in pattern recognition (fingerprints or character recognition), in stereo reconstruction (scene or aerial images), or in medical images (brain images for example). Most of these problems do not require a multimodal registration, except for some medical applications. In contrast with the profusion of the literature about registration, only a few authors report multimodal registration in medical imaging. Recent methods use mutual information in order to cope with the dissimilarity between two images [6], [7]. However, mutual information is a measure of similarity between two blocks and this approach, as well as other algorithms using block comparison, is not suitable for transformations that cannot be approximated by a translation. Another possibility is to detect features and to base the registration on these patterns.

In the case of registration of multimodal images for retinal analysis the images are particularly dissimilar. In our application area, one image is obtained after an injection of a dye, the fluorescein, in order to enhance the vascular tree of the retinal layer. The photograph is taken under ultra-violet light and therefore only injected tissues are visible. The second image is taken without any injection and is only a regular picture

Manuscript received April 6, 1998; revised March 30, 1999. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was C. Roux. Asterisk indicates corresponding author.

*F. Zana and J. C. Klein are with the Centre de Morphologie Mathématique, Ecole des Mines de Paris, 35, rue St-Honoré 77305 Fontainebleau, France.

Publisher Item Identifier S 0278-0062(99)05466-X.

filtered in the green plane. This image has a lower contrast and, compared to the fluorescein image, some patterns can have completely different gray values. Hemorrhages appear dark on both types of images, whereas vessels are white on fluorescein images and black on green images. The texture of the images is caused by different tissues and thus varies between the green images and the fluorescein images. In order to find some points for the point matching algorithm, an intermediate representation that would be common in every type of images is computed. A segmentation of the vascular tree seems the most appropriate representation to use because it maps the whole retina, does not move except in some diseases, and contains enough information for the localization of some anchor points. The algorithm for the segmentation of the vascular tree presented in [8] and [9] is then used to detect the points of bifurcation. These points, as well as the vascular tree, are features that are visible in most retinal images. We have detected them in fluorescein images and in green images and we present a technique of matching, based on this segmentation. As a consequence, we are able to register such images automatically, but since registration is based on vessel recognition, this algorithm may be extended to any other multimodal matching, provided the vascular tree can be segmented in each kind of image.

B. Registration Based on a Weak Affine Model

There exist several kinds of registration algorithms. They generally correspond to various classes of transformations—affine, stereo, weakly affine, polynomial, elastic, etc.—to various types of images—outdoor, medical, two-dimensional (2-D), three-dimensional (3-D), etc.—or to different methods of estimation—generalized Hough transform, maximal clique algorithm, graph matching, relaxation, etc. The literature in this field is so wide we will not give an exhaustive description of all possible registration algorithms here. A detailed survey of image registration techniques can be found in [10]. The present work has been inspired by the paper on fingerprints identification [11], by the paper on a registration using an iterative closest point algorithm [12] or a weak affine model [13], and by the maximal clique algorithm for stereo correspondence [14]. Its formulation also may be found quite close to [15], even though the methods used are different.

Eye fundus images are obtained either with a numerical camera or by digitizing photographs. The image distortion is caused by the eye movement as well as by the optical system of the eye and of the camera. Therefore we only see a 2-D projection of the 3-D image of the retina which should correspond to a stereo registration. However, other deformations, due to the various optical systems, may modify the accuracy of such transformations. On the other hand, restrictions can be added to the movement of the eye, whose rotation is limited, and on the scale of the camera. Our study also will be restricted to central images of the retina containing the macula, the papilla, and temporal vessels in order to limit deformations. Those simplifications reduce the projection effect caused by the optical systems. It is thus possible to

reduce the set of transformations to within a certain accuracy.

The transformations involved will be considered to be locally approximated by a composition of homothetic rotation and translation, and we will assume that the transformation is globally affine with a tolerance of ten pixels (for $512 \times 512 \times 256$ images). This value is in accordance with every image encountered and corresponds to the hypothesis of small deformations. This formulation is very well adapted to feature matching, provided that the features are distinguishable and separated by more than 20 pixels. Bifurcation points of the vascular tree are sufficiently separated to be used as basic features for the matching. The proposed algorithm can be divided into five steps:

- vessel detection and extraction of bifurcation points;
- identification of those points with invariant based on the weak affine hypothesis;
- matching process using a Hough transform algorithm and the ten-pixel tolerance;
- computation of the min-squared estimate of the global affine transformation for best scoring transformations;
- selection and more accurate computation of the most likely transformation.

C. Feature Detection in Retinal Images

Few authors report segmenting some patterns in retinal images independently of the type of the images. In [16], for the first time an algorithm for vessel detection applying Canny's matching filters [17] was tested on a green image of the retina. Another article [18] reports a method of registration of retinal images based on feature detection. Only recently [8], [9] a detection of vascular patterns was reliable enough to allow registration of retinal images provided by various sources. This detection is related to former work on crest-lines detection [19]. In [8] the cross curvature is computed and then filtered, using mathematical morphology in order to remove parts of the image that do not have a positive crosscurvature along a piecewise linear path. Positive values of the resulting filtered curvature correspond to vessel-like parts of the image. This algorithm has been tested on fluorescein images as well as green images of the retina, and the results were robust enough to consider applying this segmentation to multimodal registration.

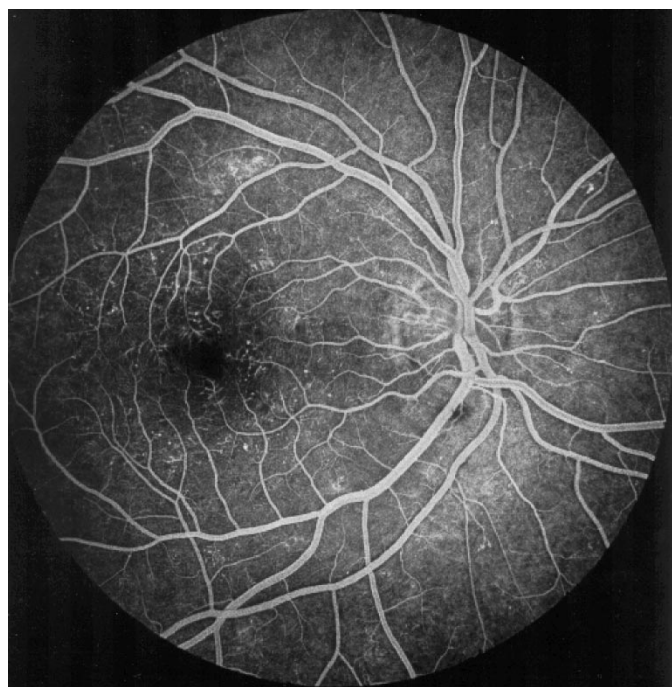
II. FEATURE RECOGNITION

A. Vessel Detection

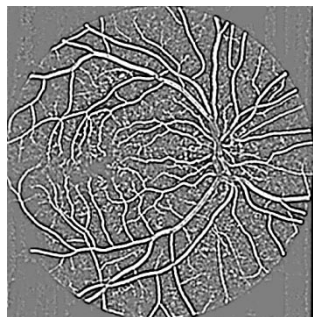
Vessels can be identified using a few characteristic properties (see Fig. 1):

- 1) they are connected;
- 2) they have a width of less than 15 pixels in $512 \times 512 \times 256$ images and are locally linear;
- 3) and in the cross direction they have a mono-modal profile.

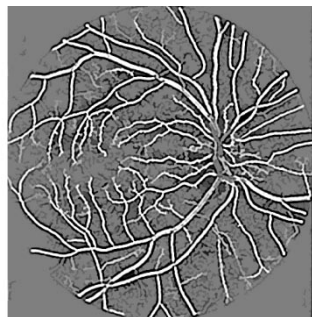
Based on the first two properties, the image is simplified by applying a supremum of opening with reconstruction using revolving linear structuring elements of size 15 pixels. It



(a)



(b)



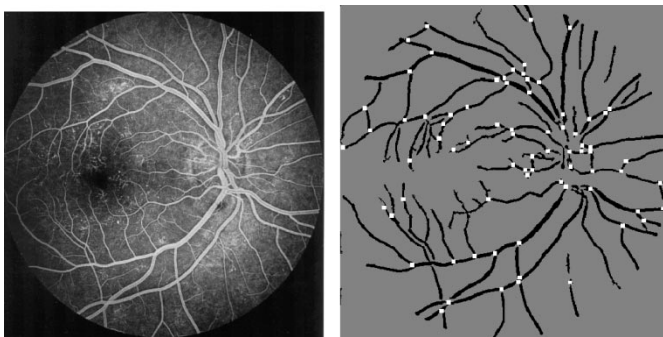
(c)

Fig. 1. Computation of the crosscurvature and noise removal. (a) Fluorescein image. (b) Crosscurvature. (c) Filtered curvature. Vessels are in white.

removes most of the noise as well as the nonlinear parts of the image. Then the contrast is enhanced, using a sum of top hats, each of which are based on an opening with one of the former linear elements. Those first modifications can be interpreted as a contrast filter for linear elements, however, other small linear patterns are recovered in this image. As explained in [8] and [9], since vessels have a linearly correlated crosscurvature, they can be differentiated from other undesirable features using a selection based on the sign of the crosscurvature and supremums of openings using linear structuring elements.

Therefore, at this stage, the crosscurvature is calculated using a principal curvature evaluation. We use a Laplacian filter as an approximation of the sign of the principal curvature as explained in [8] and [9].

Finally, the crosscurvature image is simplified, using alternating filters based on a sequence of openings and closings with linear structuring elements of increasing length (supremum of openings of size 15 with reconstruction, infimum of closings of size 15 with reconstruction, and supremum of openings of size 29).



(a)

(b)

Fig. 2. Bifurcation points.

We then detect vessel-like patterns using a simple threshold on every positive value of the filtered curvature. At this stage, fluorescein images as well as green images, have been reduced to a binary image representing the detected vascular branches. This detection can be summarized by the following steps:





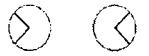


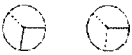

- if the image is a green image, perform an inversion;
- remove noise using an opening with reconstruction;
- sum the top-hats with linear structuring elements of various directions;
- smooth the image with a Gaussian kernel;
- compute the Laplacian;
- perform an alternating filter with linear structuring elements and reconstruction;
- Threshold positive values.

Now that each image is converted into a binary image of the vascular tree, bifurcation points are extracted and the directions of the surrounding vessels are computed in order to facilitate their recognition. More details concerning this algorithm can be found in [8] and [9]. See also [20]–[22] for deeper insight.

B. Detection of Bifurcation Points

1) *An Algorithm for the Detection of Bifurcation Points:* If the vessels were one pixel wide, a bifurcation could be detected by its T-like shape, visible in a 3×3 neighborhood. Trifurcation will be separated into two bifurcations. In order to be in such a situation, we compute a geodesic distance on the binary image of nonvascular patterns (i.e., on the inverse of the binary image of the vascular structures). Then we perform a watershed on the inverse of this distance image so that the vessels can be reduced to one-pixel-wide paths (see details of watershed implementation in [23] and [24]). Bifurcation points are detected using a supremum of openings with revolving structuring elements with a T shape. Then we examine each edge separated by two bifurcation points. When an edge is less than five pixels, or if less than 50% of this edge corresponds to a detected vessel, it is removed. Bifurcation points that are found to be close (within a ten-pixel error range) are grouped and may become a trifurcation. Direction of the surrounding edges that have not been removed are added to each bifurcation. Bifurcation points having less than two directions, or more than four directions are removed, may not be relevant, or are not robust.

TABLE I
STANDARD SITUATIONS

the vascular pattern	... detected as:
 (A)	 (a)
	 (a')
	 (b)
	 (c)
 (B)	 (d)
 (C)	 (e)

2) *The Reality of the Detection:* Apart from a false detection, the reality of the underlying vascular structure can be a little different from the detected bifurcation: the distance function can create some artifacts, introducing a bifurcation that may result only from an angle in the path followed by one vessel. We can summarize most of the cases in Table I¹ (edges in dots do not exist in the vascular structure but appear as an artifact in the distance function).

Since it is rare that the path of a vessel makes an angle such as the one in pattern (C), the corresponding detection of vessels is probably incomplete, therefore the situation (d) is usually the closest to reality. However, the left direction is certainly wrong because it is created, as an artifact, by the distance function (this direction may or may not disappear, depending on the detection of surrounding vessels). The same problem appears in case (b). The trifurcation (A) is detected as two bifurcations that are not found close enough to be merged. Case (c) corresponds to a situation where the edge joining the bifurcations has been removed, hence all directions are correct, which is not the case in (b). In order to include the local information from the directions of surrounding vessels, it is of highest importance to take into account that one direction may be completely misleading, in order to deal with false detections.

C. Model of Invariant

The detected vascular tree from which the bifurcation points are extracted can be slightly different in each image. Indeed, the following pertains.

- 1) At the early stage of diffusion of the fluorescein dye, arteries only are visible on the image.
- 2) The green image is less contrasted, therefore some smaller capillaries are missing in the corresponding binary segmentation.
- 3) An abnormality is visible in only one of the two images and covers part of the vascular tree.

¹ Every possible rotation should be considered.

As a consequence, the resulting sets of points may not match exactly, and those points that have no match in the other image risk trapping a regular algorithm in a false transformation.² Hence, an invariant should be found in order to reject obvious false matches, while taking into account the missing branches (if any) in the vascular tree. We consider I_1 and I_2 as the sets of bifurcation points in each image (including trifurcations and two-branch nodes). We assume that the transformation that transforms one point $M \in I_1$ into $M' \in I_2$ is locally approximated by a composition of translation, rotation and homothetic

$$M' = sR(\theta)M + T$$

$$R(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \quad T = \begin{pmatrix} Tx \\ Ty \end{pmatrix}. \quad (1)$$

As a consequence, angles between edges are preserved. Therefore, the selected points can be identified by the directions of the branches surrounding them by means of an appropriate measure of similarity. A maximum of four directions are associated with each point, which generate a configuration space of six dimensions (the two coordinates x and y and four directions). Points having more than four directions are rejected since they correspond neither to a bifurcation or a trifurcation.³ Each point is thus described within this space of configuration by $\underline{M} = (M, \theta_1, \theta_2, \theta_3, \theta_4)$.

For the moment, we assume that the point M matches the point M' in the sense of (1) with a rotation angle $\theta = 0$. We then consider \underline{M} and \underline{M}' and we define function d in the configuration space as follows:

$$d(\underline{M}', \theta_i) = \inf_{j=1 \dots 4, k=-1 \dots 1} \{|\theta_i - \theta'_j + 2k\pi|\}.$$

If each θ_i and each θ'_j is represented by a point on the unit circle, then $d(\underline{M}', \theta_i)$ is equal to the length of the arc from θ_i to its closest neighbor among the set of points $\{\theta'_j; j = 1 \dots 4\}$. We now symmetrize function $d(\underline{M}', \theta_i)$ with respect to both points M, M' and both series of orientations by putting

$$d(\underline{M}, \underline{M}') = \inf \left\{ \sum_{i=1}^4 d(\underline{M}', \theta_i); \sum_{j=1}^4 d(\underline{M}, \theta'_j) \right\}.$$

In addition, if M contains less than four orientations, some values are simply ignored in the corresponding sum. Function $d(\underline{M}, \underline{M}')$ turns out to be a measure of similarity of the branching context around points M and M' with regard to Table I. As a particular case, if we start from two identical configurations \underline{M} and \underline{M}' (with three directions, as in Fig. 3), and if we add one direction to configuration \underline{M}' (see Fig. 4), then the similarity measure $d(\underline{M}, \underline{M}')$ equals zero in the Euclidian case, at least. As another particular case, if \underline{M}' can be deduced from \underline{M} by a rotation of a small angle θ , then $d(\underline{M}, \underline{M}') = 4\theta$ (with four directions). This measure increases when the orientations become dissimilar. However, a weakness of this measure is the following. Starting from a set

² This fact was actually experimented using both algorithms adapted from [11] and [14].

³ Situations with more directions may be envisaged for other images, for example in the case of 3-D images.

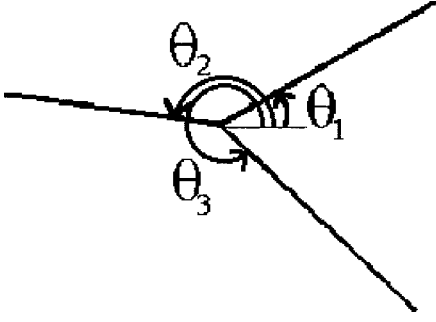


Fig. 3. Example of a bifurcation point with three directions.

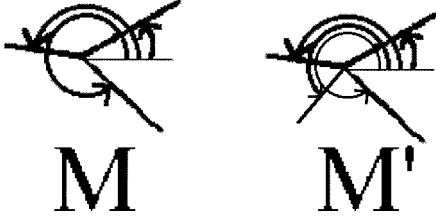
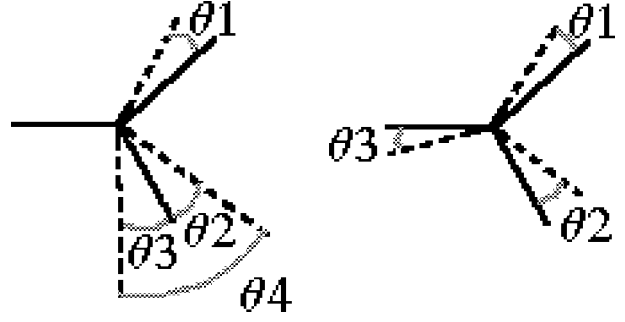
Fig. 4. Example with $d(M, M') = 0$.

TABLE II
PARAMETERS FROM FIGS. 10 AND 11. (A) REGISTRATION AS
OBTAINED IN IMAGE 10. (B) REGISTRATION AS OBTAINED IN
IMAGE 11(a). (C): REGISTRATION FAILURE, IMAGE 11(b)

$A = \begin{pmatrix} 0.954 & -0.050 \\ 0.042 & 0.956 \end{pmatrix}$	$T = \begin{pmatrix} 48.90 \\ -30.57 \end{pmatrix}$	(A)
$A = \begin{pmatrix} 0.952 & -0.070 \\ 0.043 & 0.958 \end{pmatrix}$	$T = \begin{pmatrix} 52.06 \\ -30.11 \end{pmatrix}$	(B)
$A = \begin{pmatrix} 0.861 & 0.260 \\ -0.001 & 1.047 \end{pmatrix}$	$T = \begin{pmatrix} -58.72 \\ 77.98 \end{pmatrix}$	(C)

of four directions, if we separately remove the first and third ones, for example, then the two resulting configurations will be considered as different. The measure assumes that the values of the orientations are much more reliable than the number of the orientations surrounding one point, as is precisely the case with retinal images. For example, if M and M' have three common orientations and the fourth is different, then they will be found less similar than if the fourth (and misleading) orientation of M' is missing (see Fig. 4). This property is necessary in order to handle the possible misdetections that are described in Section II-B2: Fig. 4 may correspond to case (b) in Table II.

An ambiguous case is described in Fig. 5. In the solid line, we have two different bifurcation points (with three directions each). In dotted line, we have the corresponding bifurcation points (three directions each) that are checked for matching. Both situations give the same value for $d(M, M') = \theta_1 + \theta_2 + \theta_3$. For the sake of simplicity, let us assume that $\theta_1 = \theta_2 = \theta_3$.

Fig. 5. Two examples with $d(M, M') = \theta_1 + \theta_2 + \theta_3$.

It seems obvious then that in case (b) the configurations match perfectly, but with a different rotation angle (namely, θ_1). Therefore, the value of $(\underline{M}, \underline{M}')$ reflects the dissimilarity of points conditionally to a certain rotation. If the probability of having the rotation θ_1 is very low, then those points may simply not match.

The same remark can be made concerning case (a), however, with the rotation θ_1 we have $d(M, M') = \theta_4$. Once again, if the probability of having the rotation θ_1 is low, then those points do not match. However, if the probability of θ_1 is high, under this condition the similarity measure reflects the fact that only two edges are in accordance. In fact, if this match is actually correct, the detection includes at least two problems:

The angles made by the dotted configuration are not usual, therefore it is highly probable that a vessel on the left was not detected correctly (see Table I).

If the lower left dotted angle is correctly detected, a vessel is missing on the other configuration.

These problems are emphasized in the value $d(\underline{M}, \underline{M}') = \theta_4$. As a consequence, this measure of similarity will be introduced conditionally on a transformation that has been computed based on a knowledge that does not include these directions.

In practice, the values of the orientations θ_i were computed using only the local vessel detection and were discretized using a 24-pixel circle which corresponds to an accuracy varying from 10 to 15°. Because the detection of vascular structures gives thick vessels, this discretization does not provide a fully accurate direction for each vessel. The direction of the crest line of the vessel is defined up to one pixel, therefore the measure may not necessarily be strictly zero, even when both configurations are identical. It can only be derived that two points have a certain probability for matching, deduced from their orientations and conditionally on a certain rotation. We suppose that this probability for matching depends on the former measure as follows:

$$\begin{aligned} \underline{M}_2 &= (sR(\theta)M' + T, \theta'_1 + \theta, \theta'_2 + \theta, \theta'_3 + \theta, \theta'_4 + \theta) \\ &\cdot P(M \simeq_\theta M' | s, \theta, T) = F(d(\underline{M}, \underline{M}_2)) \end{aligned} \quad (2)$$

(we note $M \simeq_\theta M'$ the event \underline{M} matches \underline{M}' with rotation θ) F can have various forms, depending on the hypotheses made about the distribution of the angles. Below, we shall take F , either null or equal to a constant value.

III. MATCHING ALGORITHM

A. A Bayesian Hough Transform

The matching algorithm is a Bayesian approach of the technique developed in [11]. We look for a global transformation that can be decomposed into a translation, a rotation and a homothetic

$$M' = sR(\theta)M + T.$$

First, each pair of couples $(M_i, M_j)(M'_k, M'_l)$ of configurations defines a unique transformation (s, θ, T) as follows:

$$\begin{aligned} s_{i,j,k,l} &= \frac{\|M'_k - M'_l\|}{\|M_i - M_j\|} \\ \theta_{i,j,k,l} &= \arg(M_i \vec{M}_j, M'_k \vec{M}'_l) \\ T_{i,j,k,l} &= \frac{1}{2}(M'_k + M'_l - sR(\theta) \cdot (M_i + M_j)). \end{aligned} \quad (3)$$

As i, j, k, l span the set of all configurations $s_{i,j,k,l}, \theta_{i,j,k,l}$, and $T_{i,j,k,l}$ generate a set of all possible transformations. Each parameter is discretized into five homogenous categories in order to facilitate computation.⁴ Then we consider the probability that the whole cluster of oriented points matches with a certain transformation. In order to estimate this probability, every pair of points is scanned in both images, the corresponding transformation is computed, and a probability for matching the whole image with this transformation is assigned to each pair $(\underline{M}, \underline{M}')$. *First step:* we test the matching between $(\underline{M}_1, \underline{M}_2)$ and $(\underline{M}'_1, \underline{M}'_2)$. This hypothesis will result in a unique transformation $(s, \theta, T) = (s_1, \theta_1, T_1)$ provided by (3) and the following probability of matching \underline{M}_1 with \underline{M}'_1 ($M_1 \sim M'_1$) and matching \underline{M}_2 with \underline{M}'_2 (i.e., $M_2 \sim M'_2$) based on (2)

$$\text{if } P(M_1 \simeq_{\theta_1} M'_1 | s_1, \theta_1, T_1) > 0 \text{ and}$$

$$P(M_2 \simeq_{\theta_1} M'_2 | s_1, \theta_1, T_1) > 0 \text{ then}$$

$$P(M_1 \sim M'_1 | s_1, \theta_1, T_1) = P(M_1 \simeq_{\theta_1} M'_1 | s_1, \theta_1, T_1)$$

$$P(M_2 \sim M'_2 | s_1, \theta_1, T_1) = P(M_2 \simeq_{\theta_1} M'_2 | s_1, \theta_1, T_1).$$

Second step: we now calculate the *a posteriori* probability of matching between both complete set of cluster ($C_M \sim C_{M'}$) under a given transformation (s, θ, T) , namely

$$P(C_M \sim C_{M'} | s, \theta, T) = 1 - \prod_i (1 - P(M_i \sim M'_i | s, \theta, T)).$$

This product holds on all pairs of points $(\underline{M}_i, \underline{M}'_i)$. Since $P(C_M \sim C_{M'})$ is a constant that is equal to one whenever it is possible to find a correct transformation, we get

$$P(s = s_o, \theta = \theta_o, T = T_o | C_M \sim C_{M'})$$

$$= P(C_M \sim C_{M'} | s = s_o, \theta = \theta_o, T = T_o) P(s_o, \theta_o, T_o).$$

$P(s_o, \theta_o, T_o)$ is the *a priori* probability of having the transformation (s_o, θ_o, T_o) that we assume to be uniformly distributed and $P(s = s_o, \theta = \theta_o, T = T_o | C_M \sim C_{M'})$ is the *a posteriori* probability based on the knowledge of the clouds

⁴These categories were found to be a good balance between the size of the space of the parameters and the complexity, in case of retinal image registration. They should be adjusted for a different problem.

of points. The transformations are then sorted with regard to their respective *a posteriori* probability. Using a functional $F(x) = 1_{x < x_o} \lambda$ with $0 < \lambda < 1$ and $N_{s,\theta,T}$, the number of matching points, the *a posteriori* probabilities are nothing but $1 - (1 - \lambda)^{N_{s,\theta,T}}$. In this case, the *a posteriori* probabilities increase with $N_{s,\theta,T}$. This number is in direct relationship with the matching score of the corresponding transformation as defined in [11], which is $N_{s,\theta,T}(N_{s,\theta,T} - 1/2)$. Such a procedure, which we can call a Bayesian Hough transformation is a direct extension of [11]. Compared with [11], only an additional storage of $P(s = s_o, \theta = \theta_o, T = T_o | C_M \sim C_{M'})$ is required. It should be noticed that the matrix in which those values are stored can be quite important, depending on the discretization of the parameters of the transformations and on the number of detected points in C_M and in $C_{M'}$.

In our case, the detection of vessels leading to the extraction of the anchor points is not reliable enough to ensure that the exact transformation will receive the highest *a posteriori* probability. Misdetected, branch occlusions, or inaccurate detection of some anchor points orient toward some exotic transformations which are totally false, but with a nonnegligible *a posteriori* probability. Therefore, the computation of these probabilities are necessary to evaluate correctly each transformation and since no uncertainty concerning the accuracy of the transformation can be tolerated, a precise selection procedure is necessary.

B. Selection of the Best Affine Estimation

1) *Affine Estimation:* Considering one possible transformation, with respect to our discretization, the best affine estimate is computed, in a min-square sense as follows:

$$A, T = \text{Arg} \min_{A, T} \sum_i P(M_i \sim M'_i | s, \theta, T) \|M'_i - AM_i - T\|^2 \quad (4)$$

with

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

The functional which is minimized can be viewed as the expectation of the quadratic error of a given affine transformation, conditioned on the best matching associated with the transformation s, θ, T . We then compute the number of points with $P(M_i \sim M'_i | s, \theta, T) > 0$ that fall in a fixed error interval (here, $\|M'_i - AM_i - T\| < 10$). If the outlier points are too numerous, then the smaller values of $P(M_i \sim M'_i | s, \theta, T)$ are put to zero and a new transformation is computed. This recursive algorithm stops because, with three points, the transformation is exact and the error is necessarily zero.

2) *Selection Based on a Minimal Error Search:* The former affine estimation (4) is computed for the 20 best *a posteriori* transformations. For each of them, the affine transformation is computed again using only those points that fall in the error interval of the previous (A, T) estimation, regardless their surrounding orientations. Such a restriction aims to achieve a maximal accuracy. A score is then computed for the transformation, based on the percentage of detected

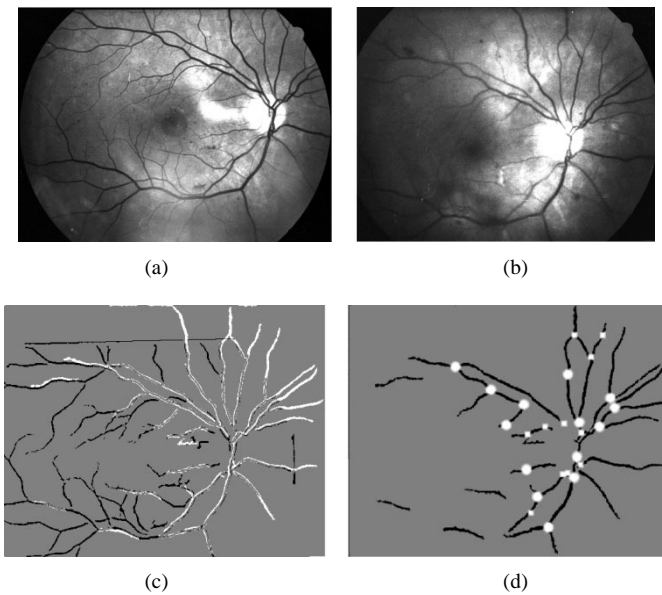


Fig. 6. Temporal registration of green images (14 anchor points). (a) Green image. (b) Green image two years later. (c) Result of the registration: comparison of the detected trees when (a) has been optimally transformed in order to fit (b). (d) Detected vessels of image (b) with bifurcation points in squared dots and anchor points in big round dots.

vessels that are superimposed. The transformation having the best score is selected. This selection procedure is necessary to ensure the robustness of the fit between both vascular structures. Since it is less time consuming than the computation of *a posteriori* probabilities, a proper balance must be reached between the accuracy of the cells of the transformations that define the size of the matrix storing the *a posteriori* probabilities and the number of the affine estimations that are to be computed to recover the good transformation.

IV. RESULTS

The former algorithm has been used to perform various kinds of registration on retinal images. A database of 17 images containing fluorescein images and green images from different patients has been used. 29 pairs of images have been tested for registration. Since the database contains two examinations of diabetic patients, spaced by two years, who were selected because they showed serious complications, it is possible to evaluate the influence of some damages that modify profoundly the detected vascular tree. In the case of diabetes, most lesions will, more or less, modify the vascular structures that we use for registration. Our database contains a patient with an important modification of its vascular tree as well as some modifications of the background texture. The bottleneck of this algorithm is the initial matching, described in Section III A, because it uses pairs of points. With m points in the first image and q points in the second, the complexity of this part is $\mathcal{O}(m^2q^2)$. However, thanks to the Bayesian formulation it is possible to compare points one by one, based on their configuration and all possible rotations. As a consequence, the registration of two $1024 \times 1024 \times 256$ images takes from 5 to 7 min on a Pentium PC 150 Mhz with 32 Mo (without MMX operations). The figures represent

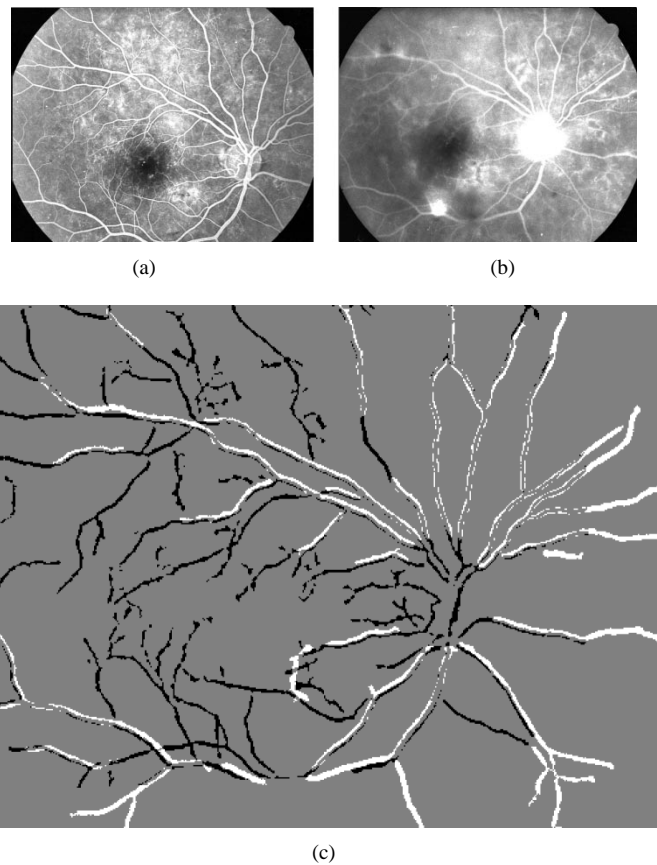


Fig. 7. Temporal registration of fluorescein images (12 anchor points). (a) Fluorescein image. (b) Fluorescein image two years later. (c) Result of the registration.

both images before registration and an image constituted of both segmentations of the vascular tree superimposed after registration.

A. Temporal Registration

Temporal registration is the first step of any algorithm that aims to measure the evolution of some lesions (for example, the turnover rate of microaneurysms [4]). Two kinds of temporal registration can be considered. Therefore the algorithm has been tested on a pair of green images and on a pair of fluorescein images. In both cases the images are spaced by two years and one can see that, in the interval, the patient showed a significant deterioration of its vascular structure. After two years, an important proportion of the smaller branches are missing, especially in the lower parts of the image, and some lesions create a moderate occlusion effect. It should be noticed that the branches of vessels in Fig. 6(b) [respectively, 7(b)] are approximately included in the vessels detected in Fig. 6(a) [respectively, 7(a)].

B. Multimodal Registration

Multimodal registration in retinal images is sometimes useful for an exact identification of the nature of the lesion. One of the most difficult pair of images corresponding to this case is presented in Fig. 8. Due to the increasing impact of diabetes, the vascular structures are rarefied. The fluorescence

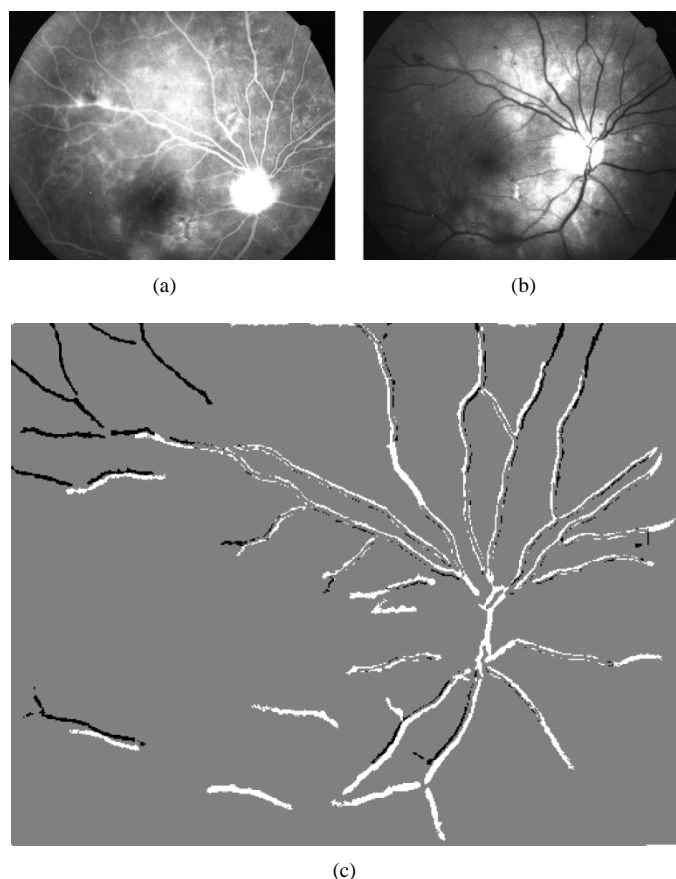


Fig. 8. Multimodal registration (13 anchor points). (a) Fluorescein image. (b) Green image during the same examination. (c) Result of the registration.

is blurred and, in this particular case, the position of the eye was an additional difficulty. Despite the great reduction of detected vessels, the quality of the registration is impressive. Although most matching points are located in the upper right corner, the lower left piece of vessel is positioned correctly within the ten-pixel approximation. However, it should be noted that both images come from the same examination, and therefore the detection of vessels is approximately the same in both images.

C. Multimodal Registration with High Differences in the Detected Vascular Tree

In order to test the behavior of the algorithm, highly dissimilar pairs of images have been processed by the algorithm. It should be noticed that the registration of both images on an intermediate fluorescein image with a complete segmented vascular tree is preferable to recover more accurately the exact registration. Therefore, those registrations do not have a particular medical interest. However, such problems are useful to evaluate the limits of the algorithm.

1) *Registration of Images Having a Nearly Complementary Segmented Vascular Tree:* In Fig. 9, the detection of the vascular tree in each image is particularly different. Because of a low concentration of the fluorescein dye in Fig. 9(a), veins are not completely detected, whereas arteries and capillaries are correctly identified. On the other hand, in Fig. 9(b) the vascular structures are deteriorated. Therefore the central capillaries are

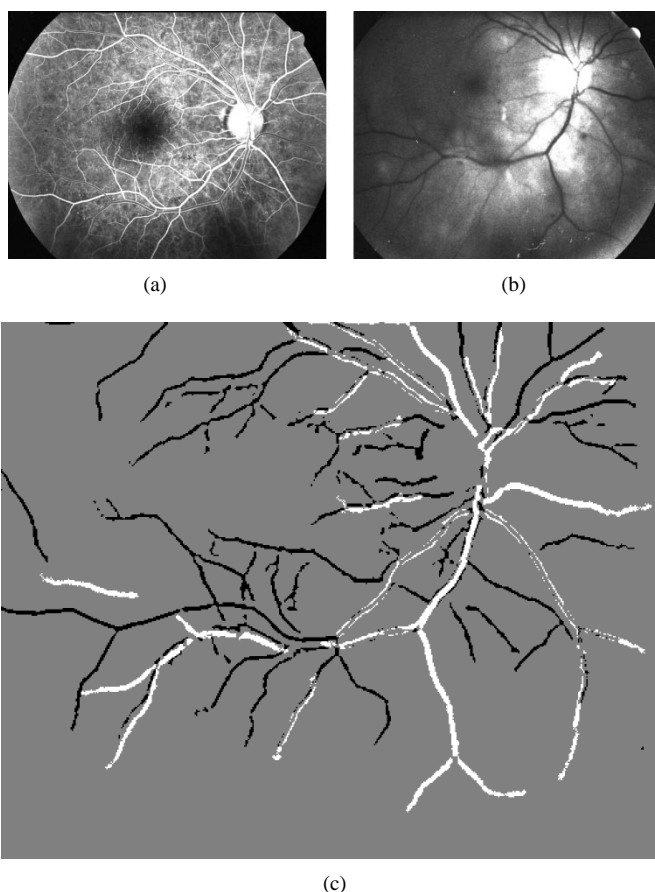


Fig. 9. Multimodal registration of images separated by two years (13 anchor points). (a) Fluorescein image. (b) Green image two years later. (c) Result of the registration.

not detected and only veins are visible. As a consequence, the common bifurcation points are not numerous and they are likely to have dissimilar surrounding orientations. The result (Fig. 9) is perfectly satisfactory.

2) *Registration with Missing Information: A Difficult Image:* In Fig. 10(a), the fluorescein angiogram is overexposed in its lower part. As a consequence, only the upper region of the vascular structure has been detected. The quality of the registration can be evaluated by the accuracy of the correspondence of the bottom part of the image. In particular, the V form in the lower central part [circled in Fig. 10(c)] of the superimposed image seems correctly registered, even though both branches of the V do not belong to the same image.

3) *Further Tests of Robustness:* In order to evaluate more precisely the robustness, we have modified image 10(a) by removing vessels and we have rerun the algorithm. Although this image is of very poor quality, several attempts were necessary in order that the registration be false. The turning point is presented in Fig. 11. In the registration Fig. 10, the chosen transformation ranks sixth, as sorted by the Bayesian Hough transform (however, the transformation obtained by the first transformation was very close, within a 1% error). Five pairs of points were selected by the Hough procedure, and all of them were found in accordance with the min-square estimation of A, T . Finally, 15 points are used to get a precise

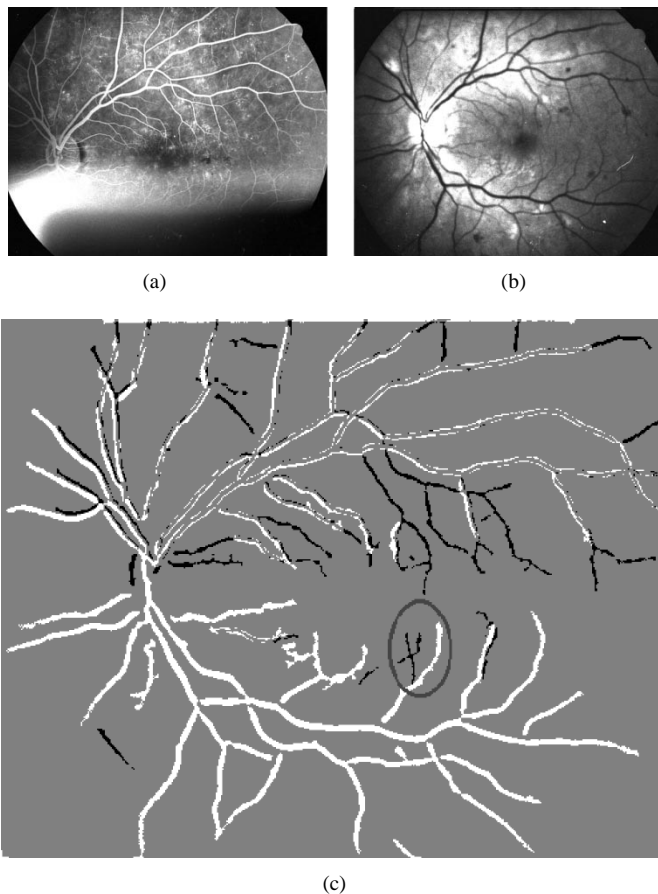


Fig. 10. Multimodal registration of images separated by two years (15 anchor points). (a) Fluorescein image. (b) Green image two years later. (c) Result of the registration.

estimation of the transformation. The resulting transformation is shown in Table II (the transformation ranked first by the Hough procedure uses ten points from the Bayesian Hough selection and 15 final points). In the registration of Fig. 11(a) with Fig. 10(b), four pairs of points were selected by the Hough procedure and one was rejected by the iterative least square fit. Finally, seven points are used to get a precise estimation of the transformation [Fig. 11(a)]. In the registration of Fig. 11(b) with Fig. 10(b), seven pairs of points were selected by the Hough procedure and only five were found correctly by the iterative least square fit. Finally, no more than five points are used to get a precise estimation of the transformation, which is wrong [Fig. 11(b)].

This example is representative of the quality of the Bayesian selection. Seventeen out of the 29 tests used the first transformation to compute the solution. Only three transformations used the iterative least square fit. These results show that the Bayesian evaluation is half responsible for most of the robustness of the method. The other half is due to the use of the best 20 transformations. The iterative least square fit in this algorithm only improves the coherence of the transformation by removing outliers.

V. CONCLUSION

An algorithm based on morphological segmentation and a Bayesian Hough transform method has been presented for

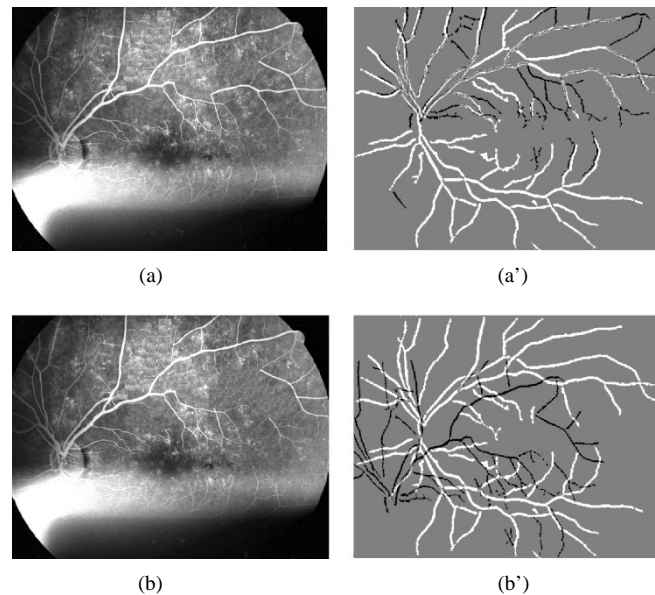


Fig. 11. The limit of failure for the algorithm: registration of Fig. 10(a) modified [(a) then (b)] with Fig. 10(b). (b') fails.

retinal image registration. This algorithm allows temporal as well as multimodal image comparison and should help the development of algorithms for retinal image analysis. Applications of such a method can be various. First of all, registration of retinal images will allow comparisons of the effect of some treatments on the eye and will help hospitals measure the correlation of alterations in different types of images. Later, it will be possible to merge this registration method with some detection algorithms, in order to facilitate or to improve the diagnosis of physicians.

ACKNOWLEDGMENT

The authors are grateful to J. Serra, Centre de Morphologie Mathématique Ecole des Mines de Paris, Paris, France, for valuable comments and corrections, as well as the referees for their interesting comments. The authors are grateful to the Eye University Hospital of Créteil, France, for providing pictures.

REFERENCES

- [1] J. Gilchrist, "Analysis of early diabetic retinopathy by computer processing of fundus images—A preliminary study," *Ophthalmic Physiol. Op.*, vol. 7, no. 4, pp. 393–399, 1987.
- [2] B. Lay, "Analyse automatique des images angiofluorographiques au cours de la rétinopathie diabétique," Ph.D. dissertation, Ecole Nat. Supérieure des Mines de Paris, Paris, France, 1983.
- [3] T. R. Friberg, J. Lace, J. Rosenstock, and P. Raskin, "Retinal microaneurysm count in diabetic retinopathy: Color photography versus fluorescein angiography," *Can. J. Ophthalm.*, vol. 22, no. 4, pp. 226–229, 1987.
- [4] M. J. Cree, J. A. Olson, K. C. McHardy, J. V. Forrester, and P. F. Sharp, "Automated microaneurysm detection," in *Proc. Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996, pp. 699–702.
- [5] F. Zana, I. Meunier, and J. C. Klein, "A region merging algorithm using mathematical morphology: Application to macula detection," *Math. Morphol. Applicat. Image Signal Processing*, vol. 1, pp. 423–430, 1998.
- [6] N. Ritter, R. Owens, K. Yogesan, and P. van Saarloos, "The application of mutual information to the registration of stereo and temporal images of the retina," in *Proc. 10th Australian Joint Conf. Artificial Intelligence: Advanced Topics in Artificial Intelligence*, vol. 1342, pp. 67–76, 1997.

- [7] W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, 1996.
- [8] F. Zana and J. C. Klein, "Robust segmentation of vessels from retinal angiography," in *Proc. Int. Conf. Digital Signal Processing*, Santorini, Greece, July 1997, pp. 1087–1091.
- [9] ———, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Trans. Image Processing*, submitted for publication.
- [10] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surveys*, vol. 24, no. 4, pp. 325–376, 1992.
- [11] N. K. Ratha, K. Karu, S. Chen, and A. K. Jain, "A real-time matching system for large fingerprint databases," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 799–812, Aug. 1996.
- [12] J. Feldmar, N. Ayache, and F. Betting, "3d-2d projective registration of free-form curves and surfaces," *Comput. Vision Image Understanding*, vol. 65, no. 3, pp. 403–424, 1997.
- [13] J. Feldmar and N. Ayache, "Rigid, affine and locally affine registration of free-form surfaces," INRIA, Research Rep. 2220, pp. 0–36, 1994.
- [14] R. Horaud and T. Skordas, "Stereo correspondence through feature grouping and maximal cliques," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 1168–1180, Nov. 1989.
- [15] W. M. Well, "Statistical approaches to feature-based object recognition," *Int. J. Comput. Vision*, vol. 21, no. 1, pp. 63–98, 1997.
- [16] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, pp. 263–269, Sept. 1989.
- [17] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 679–698, June 1986.
- [18] J. P. C. Byrne, P. G. B. Ross, P. E. Undrill, and R. P. Phillips, "Feature based retinal image registration using transporter," *Appl. Transputer*, no. 3, pp. 687–692, 1991.
- [19] R. Lengagne, P. Fua, and O. Monga, "Using crest lines to guide surface reconstruction from stereo," in *Proc. Conf. Pattern Recognition*, Aug. 1996.
- [20] J. Serra, *Image Analysis and Mathematical Morphology*. London, U.K.: Academic, 1982.
- [21] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. R. Soc. London*, vol. 207, pp. 187–217, 1980.
- [22] P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen, "A survey of thresholding techniques," *Comput. Vision, Graph. Image Processing*, vol. 41, pp. 233–260, 1988.
- [23] S. Beucher and C. Lantuejoul, "The use of watershed in contour detection," in *Proc. Int. Workshop Image Processing, Realtime Edge Motion Detection/Estimation*, 1979.
- [24] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," *Math. Morphol. Image Processing*, 1992, pp. 433–481.
- [25] M. P. Do Carmo, *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ: Prentice-Hall, 1976.
- [26] W. Hoff and N. Ahuja, "Surface from stereo: Integrating feature matching, disparity estimation, and contour detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 121–136, Feb. 1989.
- [27] H.-C. Liu and M. D. Srinath, "Partial shape classification using contour matching in distance transformation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 1072–1079, Nov. 1990.
- [28] O. Monga, N. Armande, and P. Montesinos, "Thin nets and crest lines: Application on satellite data and medical images," INRIA, Research Rep. 2480, Feb. 1995.