Quantitative evaluation of image registration techniques in the case of retinal images.

Yann Gavet, Mathieu Fernandes and Jean-Charles Pinoli

École des Mines de Saint-Étienne,

CIS-LPMG/CNRS,

158 Cours Fauriel, 42023 SAINT-ÉTIENNE, FRANCE

To cite this article, use:

@ARTICLE{JEIGavet2012,

}

```
author = {Yann Gavet and Mathieu Fernandes and Jean-Charles Pinoli},
```

```
title = {Quantitative evaluation of image registration techniques in the case of retinal images.},
journal = {Journal of Electronic Imaging},
year = {2012},
doi = {10.1117/1.JEI.21.2.021118},
volume = {21},
pages = {021118},
number = {2}
```

ABSTRACT

In human retina observation (with non mydriatic optical microscopes), an image registration process is often employed to enlarge the field of view. Analyzing all the images takes a lot of time. Numerous techniques have been proposed to perform the registration process. Its good evaluation is a difficult question that is then raising.

This article presents the use of two quantitative criterions to evaluate and compare some classical featurebased image registration techniques. The images are first segmented and the resulting binary images are then registered. The good quality of the registration process is evaluated with a normalized criterion based on the ϵ dissimilarity criterion, and the Figure of merit criterion (*fom*), for 25 pairs of images with a manual selection of control points. These criterions are normalized by the results of the affine method (considered as the most

Further author information: (Send correspondence to Y. Gavet)

Y. Gavet: E-mail: gavet@emse.fr, Telephone: [33] 4 77 42 01 70

simple method). Then, for each pair, the influence of the number of points used to perform the registration is evaluated.

Keywords: Image registration, Human retina, Dissimilarity criterion

1. INTRODUCTION

An image registration method consists in representing two or more images into the same spatial coordinate system, such that there is a correspondence between the structures of the images. For more information, a survey is presented in [1]. These different methods have different efficiency levels, which can be evaluated. A visual assessment can be performed, but there is not quantification. Some similarity measures can be used [2–4]. This article proposes the use of a novel dissimilarity criterion [5] to evaluate quantitatively the performance of different image registration methods. It will also make use of the figure of merit [6] to compare the results to a more established criterion. This article is focused on these criterions and is not intended to be a survey on registration methods. The problem of the quality control on retinal images registration is addressed as a relevant illustration, but the proposed dissimilarity criterion can be employed in a lot of different application areas such as the the control of image segmentation processes.

Image registration techniques can be divided into two categories: intensity-based (iconic) registration and feature-based registration techniques. In the case of retinas, iconic registration is not a good choice because images present different illumination conditions and are generally textureless except in the neighborhood of the vessels. Moreover, one can notice that the ophthalmologists manually perform this registration process by using a landmark matching based on the vessels and their bifurcations [7]. A fully automated registration technique on the retina images would segment the vessels, find the control points (i.e. the bifurcations) and apply a matching algorithm between the different images.

1.1 The ϵ dissimilarity criterion

The recently reported ϵ dissimilarity criterion [5] is defined as a comparison either between continuous sets or between discrete sets. In this article, the considered sets are discrete, e.g. finite sets of pixels. Practically, these sets can be the results of some image segmentation process. This novel criterion introduces a tolerance in the measure (thus, it is not mathematically a metric) by the mean of the Minkowski addition [8]. The reader can refer to [9] to get more information on evaluation of human retina images segmentations.

1.1.1 Minkowski addition

The Minkowski addition [8] defines an algebraic operation between sets in \mathbb{R}^n .

If M and B are two sets of \mathbb{R}^n , the Minkowski sum of M and B, denoted $M \oplus B$, is then defined by Eq. 1:

$$M \oplus B = \{m + b | m \in M, b \in B\}$$
$$= \bigcup_{b \in B} \{m + b | m \in M\}$$
(1)

where \oplus is the so-called Minkowski addition symbol. In the field of Mathematical Morphology [10], it corresponds to the dilation, and *B* is called a structuring element (for example a ball).

1.1.2 Definition of the ϵ dissimilarity criterion

The ϵ dissimilarity criterion with the tolerance ρ applied to segmented (binary) images is defined in the case of discrete (numerical) images by the following equation [5]:

$$\epsilon_M^{\rho}(X) = \frac{\#\{(X \setminus M \oplus \rho N) \cup (M \setminus X \oplus \rho N)\}}{\#\{M \oplus \rho N\}}$$
(2)

with N being the structuring element of radius 1 (the unit ball) and # denoting the number of pixels in the set (# is the cardinal operator, counting the number of non null pixels in the sets X or M. Roughly speaking, the reader can see this formula as a tolerant and normalized version of the symmetric difference). Practically, the tolerance parameter ρ is the radius of the ball used to dilate the binary images. The choice of this value is not the purpose of this article, though it remains an important question. The tolerance can be illustrated by Fig. 3, where some vessels that are really close would be considered as different in the case no tolerance would be used. As shown in [5], the ϵ dissimilarity criterion is tolerant to small variations, like thickness variations or translations, that are visually not important. Moreover, it is tolerant to punctual problems, whereas (among others) the figure of merit is not.

1.2 Figure of merit

The figure of merit (attributed to Pratt, [6, 11]) is a classical metric that can be used to compare image segmentations. It is defined by Eq 3:

$$f_M(X) = 1 - \frac{1}{\max\{\#(M), \#(X)\}} \sum_{p \in X} \frac{1}{1 + d^2(p, M)}$$
(3)

where d(p, M) is the Euclidean distance between the pixel $p \in X$ to the closest pixel of M and # the number of pixels of the considered segmentation M or X (which are non empty images).

1.3 Structure of this article

The next section will present the retina image database used to perform the tests, the process used to segment the vessels and the different registration methods. Then, in the following section (3), the quantitative results will be presented. The accuracies of the registration methods will be compared together. The influence of the number of points in this accuracy will also be studied.

2. METHOD

2.1 Human retina image database

A database of 25 pairs of retina images has been constructed. For each pair of images, the points of correspondence (called control points) have been manually marked by an expert (with a minimum of 15 points of correspondence for each pair). Generally, these points are automatically extracted [12], but this remains a difficult task. The manual selection of the pairs of points thus avoids the difficult problem of their automatic detection (see Fig. 1) and their automatic pairing. The minimal number of control points to use for each registration technique is presented in table 1.

Geometric transformation / Registration method	Minimal number of control points		
Affine	3		
Similitude	5		
Projective	3		
Polynomial of order 2	6		
Polynomial of order 3	10		
Piecewise linear	4		
Local linear	2		

Table 1. Minimal number of control points for each image registration method.

Additionally, the (almost circular) region of interest of the images are extracted (see Fig. 2): this will be used as a mask to constrain the overlapping areas.

2.2 Image segmentation

A segmentation process extracts the vessels present in the images. A lot of different methods exists for achieving this task [13–16]. In this article, only one method is employed to get the vessels: the one of Chanwimaluang [15]. Two examples are presented in Fig. 2(c) and 2(d). These segmented vessels will be used later for the quantitative evaluation of some registration techniques.

As we choose to evaluate the registration process by the evaluation of the registration of the segmentation, it is important to notice that some other segmentation evaluation criterions could have also be used (see [5] for a detailed study). The main drawback of these methods is that they are not tolerant to small spatial variations of the segmented objects.

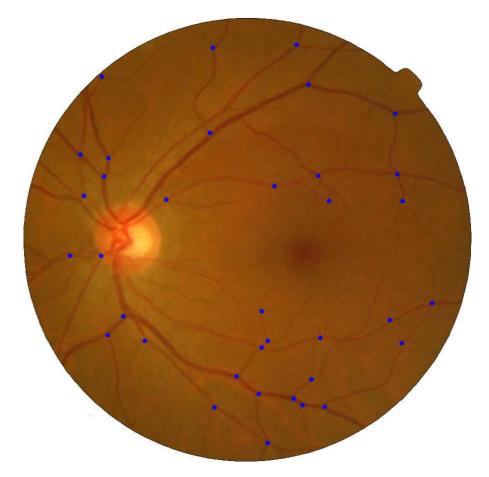
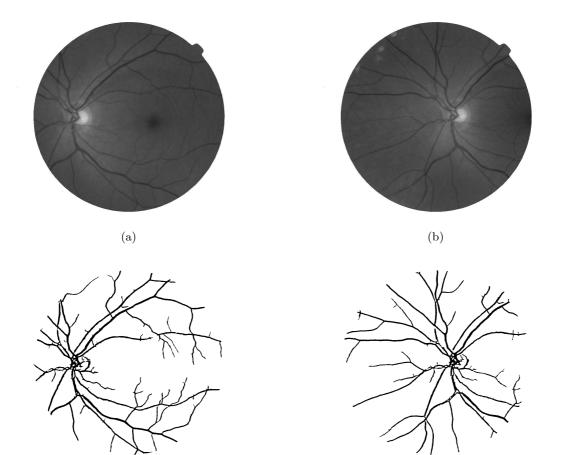


Figure 1. Example of a human retina image with the manual selection of control points (the dark squares). This selection is performed for each pair of images, avoiding to automatic pairing problem. The optical nerve appears as the bright disk. The dark area in the center of the image is the fovea, the optical center of the retina.



(c)
(d)
Figure 2. Example of a pair of human retina images. The represented images (a) and (b) are the green component of the RGB images. The binary result of the segmentation (see subsec. 2.2) of these images is respectively represented in (c) and (d) for the images (a) and (b).

2.3 Image registration

The registration process consists in finding a geometric spatial transformation to be applied to one image to fit the other. Thus, the computation of this transformation is performed on the sets of corresponding points. If u and v are the coordinates of a point in the image U (as Unregistered) to be registered, and x and y its coordinates in the second image B (as **B**ase image), the geometric transformation used belongs to one of the following categories presented hereafter (see [1, 17] for more details on these methods):

• Similitude: the geometric transformation is a similitude (a combination of translation (h and k), rotation θ and scaling s).

$$u = x \cdot s \cos \theta - y \cdot s \sin \theta + h \tag{4}$$

$$v = x \cdot s \sin \theta + y \cdot s \cos \theta + k \tag{5}$$

• Affine: a first order polynomial relation between the two systems of coordinates.

$$u = ax + by + c \tag{6}$$

$$v = dx + ey + f \tag{7}$$

• Projective: this geometric transformation may be used in the case of nonlinearities in the lens or sensor.

$$u = \frac{ax+by+c}{dx+ey+1} \tag{8}$$

$$v = \frac{fx + gy + h}{dx + ey + 1} \tag{9}$$

• Polynomial: the relation between the coordinates is polynomial. Let P and Q be two polynomials for which the maximum degree in x and y is n. Practically, n is 2 or 3. If n = 1, the geometric transformation is affine.

$$u = P(x, y) \tag{10}$$

$$v = Q(x, y) \tag{11}$$

- Piecewise linear: a linear transformation function is used to map a triangle (defined by a Delaunay triangulation of the control points, see [18]) in the registered image to the corresponding triangle in the base image.
- Local linear: Fernandes et *al.* [12] proposed a method that computes a transformation for each point in the images from its two nearest control points.

3. QUANTITATIVE EVALUATION AND RESULTS

The aim of this article is to present an application case where the ϵ dissimilarity criterion can be used. This can be seen as a test result for the theoretical information that can be found in [5]. It is focused on the comparison of registration processes.

3.1 Mean results over the database

The registration process is performed on the segmented images of human retinas. Only the overlapping regions are kept (with the use of a mask). The dissimilarity criterions ϵ and f are applied on them (the value $\rho = 3$ is chosen). To compare the different registration methods together, the results are normalized by the results of the affine method, which is considered as a reference. The reason of this choice is because this is the simplest registration method and this is what a human would do with printed photographs. Notice that the lower the criterion value is, the better the registration method is.

Let *B* be the base (segmented) image, and *U* the unregistered (segmented) image. The transformation *T* maps *U* into the coordinate system of *B*. If *A* is the affine transformation, the normalized ϵ criterion, denoted $\tilde{\epsilon}$ (Eq. 12), and the normalized figure of merit, denoted \tilde{f} (Eq. 13), are computed as follows:

$$\tilde{\epsilon}_B(T(U)) = \frac{\epsilon_B^3(T(U))}{\epsilon_B^3(A(U))}$$
(12)

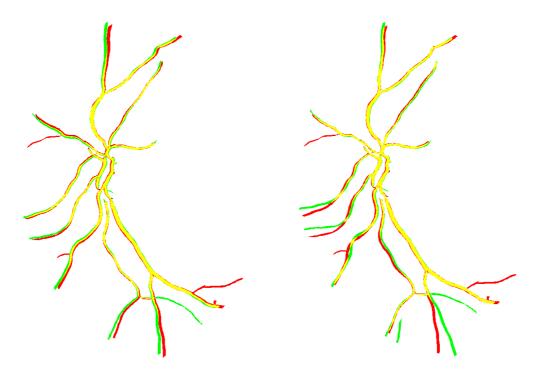
$$\tilde{f}_B(T(U)) = \frac{f_B(T(U))}{f_B(A(U))}$$
(13)

- $\tilde{\epsilon} \leq 1$ ($\tilde{f} \leq 1$) means that the geometric transformation gives a better result than the affine transformation, according to the given criterion.
- $\tilde{\epsilon} \ge 1$ ($\tilde{f} \ge 1$) means that the geometric transformation gives a lower result than the affine transformation, according to the given criterion.

The results are proposed in Fig. 4 and Tabs. 3, 4 and 2. They show that the second order polynomial method appears better than the others, probably due to the original curved (almost spherical) shape of human retinas. The local linear method would be a good choice, but presents high values in some cases.

About the comparison criterions, they give almost the same results, but the $\tilde{\epsilon}$ dissimilarity criterion gives more precise information, as expected. Take for example pair 7 (see Fig. 3). The $\tilde{\epsilon}$ dissimilarity criterion clearly distinguishes the two methods (1 versus 1.19, see Tab. 3), whereas the \tilde{f} criterion gets the same result (1, see Tab. 4).

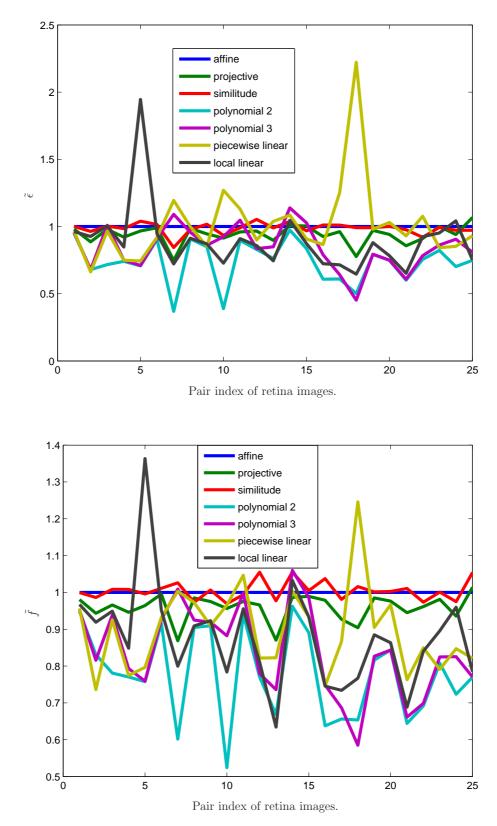
These results must be tempered by the fact that they depend on the choice of the control points. This is why the next subsection will try to evaluate their influence.



(a) Result of the registration process of pair 7 with the affine method.

(b) Result of the registration process of pair 7 with the piecewise linear method.

Figure 3. These figures illustrate the registration processes for the affine and piecewise linear methods. The images are restrained to the overlapping areas. The base image is represented by the red channel, the registered image is represented by the green channel. The superimposition appears in yellow. The $\tilde{\epsilon}$ and \tilde{f} criterions give contradictory results: for $\tilde{\epsilon}$, the affine method gives a better result than the piecewise linear method (1 compared to 1.19, see Tab. 3), but for \tilde{f} , affine and piecewise linear methods are equivalent (see Tab. 4). Visually, our opinion tends to prefer the result presented in (a) like $\tilde{\epsilon}$ does. This means that the figure of merit (by the use of \tilde{f}) is cannot be trusted in making a comparison between these image registration methods.



(b) Evolution of criterion \tilde{f} over all images of the database.

Figure 4. Quantitative comparison between the 8 image registration methods using the ϵ dissimilarity criterion (a) and the figure of merit \tilde{f} (b). These graphs represent the values of the criterions (normalized by the affine registration method) for 25 images of the human retina database. The second order polynomial method gives the best results in average for both criterions.

Geometric Transformation / Registration method	Mean $\tilde{\epsilon}$	Mean \tilde{f}
Affine	1	1
Similitude	0.94	0.96
Projective	0.99	1
Polynomial of order 2	0.74	0.78
Polynomial of order 3	0.85	0.84
Piecewise linear	1.01	0.90
Local linear	0.89	0.88

Table 2. Results over the 25 images of the human retina database. The best average value is obtained by the second order polynomial method. The local linear method gives generally good results, but sometimes the distortion of the registered images is too strong (see Fig. 4).

3.2 Influence of the number of control points

To evaluate the influence of the number of control points, one pair of corresponding human retina images is taken (Pair number 25). This pair of images contains 22 pairs of control points. This original result is a practical answer to the choice of the number of control points. One would say the larger, the better. In fact, this is not always the truth.

The results are presented in Fig. 5. The number of control points is randomly chosen (among the 22 points) and the average ϵ value is represented is abscissa of 100 random choices. $\epsilon_B^3(T(U))$ (measuring the ϵ dissimilarity of the transformation of the unregistered image T(U) compared to the base image B) is represented in ordinates (Fig. 5(a)). The same graph for the *fom* criterion is presented in Fig. 5(b). Notice that each method requires a minimum number of pairs of control points.

It appears that the local linear method gives interesting results. For all methods but the piecewise linear method, the more number of control points are chosen, the better it performs. For this latter, the spreading out of the control points over the images more strongly impacts the results.

4. CONCLUSION AND PERSPECTIVES

The problem of retinal images registration is adressed as an illustration to the use of the ϵ dissimilarity criterion. In this article, some classical registration techniques are evaluated and compared. It is shown that the ϵ criterion is fairly simple to understand and manipulate. These results have also shown that the ϵ dissimilarity criterion, with its tolerance, is making a discrimination between results that the figure or merit cannot distinguish. As

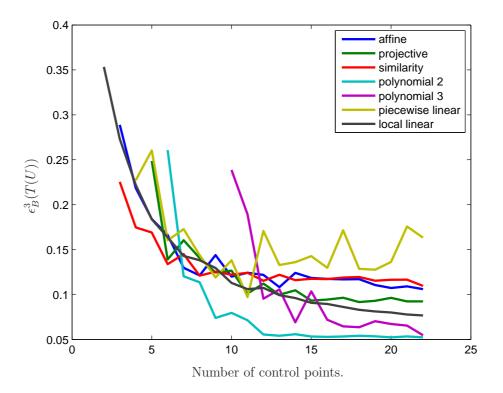
	affine	projective	similitude	polynomial	polynomial	piecewise	local linear
				2	3	linear	
Pair 1	1.00	0.98	1.00	0.95	0.96	0.97	0.96
Pair 2	1.00	0.89	0.96	0.68	0.68	0.66	0.93
Pair 3	1.00	0.98	1.00	0.72	1.01	0.97	1.01
Pair 4	1.00	0.92	0.98	0.74	0.75	0.75	0.85
Pair 5	1.00	0.97	1.04	0.71	0.71	0.74	1.95
Pair 6	1.00	0.99	1.01	0.90	0.91	0.92	0.95
Pair 7	1.00	0.75	0.84	0.37	1.09	1.19	0.72
Pair 8	1.00	0.99	0.98	0.91	0.96	1.00	0.91
Pair 9	1.00	0.95	1.02	0.85	0.86	0.85	0.87
Pair 10	1.00	0.91	0.93	0.39	0.92	1.27	0.73
Pair 11	1.00	0.96	0.99	0.90	1.05	1.13	0.91
Pair 12	1.00	0.97	1.05	0.83	0.84	0.90	0.86
Pair 13	1.00	0.90	0.99	0.76	0.85	1.04	0.75
Pair 14	1.00	1.01	1.04	0.97	1.14	1.08	1.05
Pair 15	1.00	1.00	0.97	0.84	1.03	0.91	0.87
Pair 16	1.00	0.93	1.01	0.61	0.79	0.87	0.72
Pair 17	1.00	0.96	1.01	0.61	0.64	1.25	0.72
Pair 18	1.00	0.78	0.99	0.50	0.45	2.22	0.65
Pair 19	1.00	0.97	0.99	0.79	0.79	0.98	0.88
Pair 20	1.00	0.94	1.00	0.75	0.75	1.03	0.78
Pair 21	1.00	0.86	0.98	0.60	0.61	0.93	0.65
Pair 22	1.00	0.91	0.92	0.76	0.78	1.08	0.92
Pair 23	1.00	1.00	0.99	0.82	0.86	0.84	0.95
Pair 24	1.00	0.94	0.97	0.70	0.91	0.85	1.04
Pair 25	1.00	1.07	0.97	0.75	0.81	0.93	0.75

Table 3. This table presents the normalized results ($\tilde{\epsilon}$) for the 25 pairs of retina images and for the different image

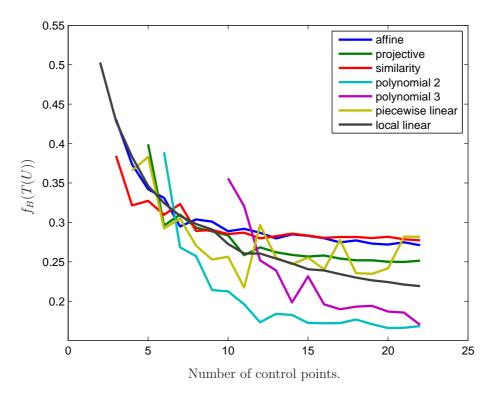
registration methods. The lower values show the best results. The first column presents the number of the pair of retina images in the database.

	affine	projective	similitude	polynomial	polynomial	piecewise	local linear
				2	3	linear	
Pair 1	1.00	0.98	1.00	0.95	0.95	0.96	0.97
Pair 2	1.00	0.94	0.99	0.83	0.82	0.74	0.92
Pair 3	1.00	0.97	1.01	0.78	0.94	0.93	0.95
Pair 4	1.00	0.95	1.01	0.77	0.79	0.77	0.85
Pair 5	1.00	0.96	1.00	0.76	0.76	0.80	1.36
Pair 6	1.00	1.00	1.01	0.92	0.93	0.93	0.95
Pair 7	1.00	0.87	1.03	0.60	1.01	1.00	0.80
Pair 8	1.00	0.99	0.97	0.90	0.92	0.97	0.91
Pair 9	1.00	0.98	1.01	0.91	0.92	0.91	0.92
Pair 10	1.00	0.96	0.97	0.52	0.88	0.97	0.78
Pair 11	1.00	0.98	0.99	0.94	1.00	1.05	0.96
Pair 12	1.00	0.97	1.05	0.77	0.78	0.82	0.81
Pair 13	1.00	0.87	0.98	0.67	0.74	0.82	0.63
Pair 14	1.00	0.99	1.06	0.96	1.06	1.00	1.03
Pair 15	1.00	0.99	1.00	0.89	0.99	0.93	0.93
Pair 16	1.00	0.98	1.04	0.64	0.75	0.75	0.75
Pair 17	1.00	0.93	0.98	0.66	0.69	0.87	0.73
Pair 18	1.00	0.90	1.02	0.65	0.59	1.25	0.77
Pair 19	1.00	0.99	1.00	0.82	0.83	0.90	0.88
Pair 20	1.00	0.98	1.00	0.84	0.84	0.97	0.86
Pair 21	1.00	0.94	1.01	0.64	0.66	0.76	0.69
Pair 22	1.00	0.96	0.97	0.69	0.70	0.85	0.84
Pair 23	1.00	0.98	1.00	0.81	0.82	0.79	0.90
Pair 24	1.00	0.94	0.97	0.72	0.83	0.85	0.96
Pair 25	1.00	1.01	1.05	0.77	0.77	0.82	0.78

Table 4. This table presents the normalized results (\tilde{f}) for the 25 pairs of retina images and for the different image registration methods, computed with the fom criterion. The lower values show the best results. The first column presents the number of the pair of retina images in the database.



(a) Evolution of the ϵ criterion while using a certain number of control points.



(b) Evolution of the *fom* criterion while using a certain number of control points. Figure 5. This figure shows the evaluation of the registration in regard to the number of control points. This example is computed for only one pair of retina images. It shows that the so-called local linear method generally performs well even when a small number of points is used. Notice that ϵ and *fom* give almost the same results.

expected, the ϵ dissimilarity criterion is adapted to compare objects (shapes, sets, segmentations...) in a visual way, i.e. in a manner the visual perception system behaves.

The reader can notice that ϵ could also be employed to evaluate results of segmentation, which was the reason why it was originally introduced [5].

ACKNOWLEDGMENTS

The authors wish to thank Pr. Gilles Thuret and Pr. Philippe Gain from University Hospital Centre, Saint-Etienne, France for supporting this work and for providing pictures of human retinas.

REFERENCES

- Zitova, B., "Image registration methods: a survey," Image and Vision Computing 21, 977–1000 (October 2003).
- [2] Ritter, N., Owens, R., Cooper, J., Eikelboom, R., and Van Saarloos, P., "Registration of stereo and temporal images of the retina," *Medical Imaging, IEEE Transactions on* 18, 404–418 (may 1999).
- [3] Skerl, D., Likar, B., and Pernus, F., "A protocol for evaluation of similarity measures for rigid registration," *Medical Imaging, IEEE Transactions on* 25, 779–791 (june 2006).
- [4] Lalibert, F., Gagnon, L., and Sheng, Y., "Registration and fusion of retinal images-an evaluation study," Medical Imaging, IEEE Transactions on 22, 661–673 (May 2003).
- [5] Gavet, Y. and Pinoli, J.-C., "A geometric dissimilarity criterion between jordan spatial mosaics. theoretical aspects and application to segmentation evaluation," *Journal of Mathematical Imaging and Vision* 42, 25–49 (2012). 10.1007/s10851-011-0272-4.
- [6] Abdou, I. and Pratt, W., "Qualitative design and evaluation of enhancement/thresholding edge detector," *Proc. IEEE.* 67, 753–763 (May 1979).
- [7] Can, A., Stewart, C. V., Roysam, B., and Tanenbaum, H. L., "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Transactions on Pattern Analysis* and Machine Intelligence 24, 347–364 (2002).
- [8] Minkowski, H., "Volumen und Oberfläche," Mathematische Annalen 57, 447–495 (1903).
- [9] Gegundez, M., Aquino, A., Bravo, J., and Marin, D., "A function for quality evaluation of retinal vessel segmentations," *Medical Imaging, IEEE Transactions on* **31**(2), 231–239 (2012).
- [10] Soille, P., [Morphological Image Analysis: Principles and Applications], Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2nd ed. (2003).

- [11] Strasters, K. C. and Gerbrands, J. J., "Three-dimensional image segmentation using a split, merge and group approach," *Pattern Recognition Letters* 12(5), 307–325 (1991).
- [12] Fernandes, M., Gavet, Y., and Pinoli, J.-C., "A feature-based dense local registration of pairs of retinal images.," in [Proceedings of the 4th International Conference on Computer VISion Theory and APplications, VISAPP 2009 ISBN=978-989-8111-69-2 VISAPP 2009: 4th International Conference on Computer VISion Theory and APplications], Ranchordas, A. ; Araujo, H., ed., 1, 265–265, INSTICC-Institut for Systems and Technologies of Information Control and Communication, Lisbonne Portugal (2009).
- [13] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., and Goldbaum, M., "Detection of blood vessels in retinal images using two-dimensional matched filters," *Medical Imaging, IEEE Transactions on* 8, 263–269 (sep 1989).
- [14] Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., and van Ginneken, B., "Ridge-based vessel segmentation in color images of the retina," *Medical Imaging, IEEE Transactions on* 23, 501–509 (april 2004).
- [15] Chanwimaluang, T., Fan, G., and Fransen, S. R., "Hybrid retinal image registration," *IEEE Transactions on Information Technology in Biomedicine* 10(1), 129–142 (2006).
- [16] Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., and Abramoff, M., "Comparative study of retinal vessel segmentation methods on a new publicly available database," in [SPIE Medical Imaging], Fitzpatrick, J. M. and Sonka, M., eds., 5370, 648–656, SPIE, SPIE (2004).
- [17] IMGFSR, "Transformation functions for image registration," tech. rep., Image Fusion Systems Research.
- [18] Goshtasby, A., "Piecewise linear mapping functions for image registration," *Pattern Recognition* 19(6), 459–466 (1986).