Comparison of Different Global and Local Automatic Registration Schemes: An Application to Retinal Images

Evangelia Karali, Pantelis Asvestas, Konstantina S. Nikita, and George K. Matsopoulos

National Technical University of Athens, School of Electrical and Computer Engineering, Iroon Politechniou 9, Zografos 15780, Greece
ekarali@biosim.ntua.gr

Abstract. In this paper, different global and local automatic registration schemes are compared in terms of accuracy and efficiency. The accuracy of different optimization strategies based on a variety of similarity measures (cross-correlation, mutual information coefficient or chamfer distance) is assessed by means of statistical tests. Results from every optimization procedure are quantitatively evaluated with respect to the gold-standard (manual) registration. The comparison has shown that chamfer distance is a robust and fast similarity measure that can be successfully combined with common optimization techniques in retinal image registration applications.

1 Introduction

Retinal images are the common diagnostic tool in ophthalmology. Many eye diseases, like diabetic retinopathy, glaucoma, cataract and age-related macular degeneration can be detected in fundus images as well as many therapeutic techniques are planned and implemented according to eye vessels topography, as it is presented in ophthalmic images [1]. Comparison studies of ophthalmic images require thorough visual inspection because of their spatial misalignment, due to changes in the geometry between fundus camera and the retina or changes in retinal vessel topography because of pathological conditions, like glaucoma. Manual registration is the standard method used in clinical practice, however it depends on human knowledge and experience [2]. On the other hand, many automatic registration schemes that combine speed and accuracy have been applied to retinal images [1][3].

The most common ophthalmic imaging techniques are fluorescein angiography (FA) and indocyannine green angiography (ICG). FA images the fluorescence of a dye, fluoroscein, as it travels through retinal vessels because of blood circulation. Soon after intravenous injection to the patient of sodium fluoroscein 10% (usually after 5-7sec), FA images are obtained at a rate of 1 image/sec for the next 20sec. Prior to any examination, a Red-Free (RF) retinal image is acquired using a bandpass green filter, which cuts of the red light. In RF images, retinal blood vessels appear dark. Information from RF images in combination with FA and/or ICG data is used for the evaluation of disease progress [3].

C. Barillot, D.R. Haynor, and P. Hellier (Eds.): MICCAI 2004, LNCS 3216, pp. 813–820, 2004.

In this work, automatic registration schemes based on various optimization techniques and on intrinsic image characteristics are compared in terms of accuracy and efficiency. In particular, three standard similarity measures, namely cross-correlation coefficient (C_{cc}), mutual information coefficient (MI) and chamfer distance (CD) have been used in combination with four different common optimization algorithms: Downhill Simplex (DSM), Powell's Method (PM), their combination (DSM-PM) and a combination of Simulated Annealing (SA) with PM (SA-PM). The accuracy of the different registration schemes has been assessed by means of statistical tests. Results from every optimization procedure have been quantitatively evaluated with respect to the gold-standard (manual) registration.

2 Materials and Methods

2.1 Image Acquisition

Retinal images were obtained using the IMAGEnet 1024 system, a fundus camera that provides 50% of coverage, 39mm working distance and specials filters for FA and acquired digital ophthalmic images 1024x1024 pixels in size. The automatic and manual registration techniques were applied to retinal images 512x512 pixels in size to increase optimization algorithm convergence speed.

2.2 Image Preprocessing

No preprocessing step was required for registration schemes based on C_{cc} or MI. Optimization techniques based on minimization of CD were applied to edge images of the retina, which were derived from the corresponding gray level images by applying first a canny edge detector with standard deviation $\sigma = 3$ and then the reconstruction opening operator that links edge fragments [4].

2.3 Registration Schemes

Every registration method is determined by the chosen transformation model, the similarity measure and the optimization strategy [5].

Transformation Model. The most suitable transformation model for registering retinal image pairs is a two dimensional (2D) affine transformation [5] that maps every pixel (x, y) of an image I to a pixel (x', y') of a reference image J according to the equation:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \end{pmatrix} . \quad (1)$$

Similarity Measures. *Cross-correlation coefficient* (C_{cc}) is suitable for registering monomodal medical images [1]. The C_{cc} between two images *I* and J *M*x*N* pixels in size is mathematically expressed by:

$$C_{cc} = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(I(x,y)-\bar{I}\right)\left(J(x,y)-\bar{J}\right)}{\sqrt{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(I(x,y)-\bar{I}\right)^2\sum_{x=1}^{M}\sum_{y=1}^{N}\left(J(x,y)-\bar{J}\right)^2}} \cdot \quad (2)$$

where $\bar{I}$ and $\bar{J}$ are the mean gray values of *I* and *J* respectively and $C_{cc} \in$ [-1.1].

Mutual Information (MI) can be considered as a generalized non-linear correlation function. Considering two images *I* and *J*, which are geometrically associated by a transformation T, then if *a* and *b* are the gray values of *I*(*x*,*y*) and *J*(T(*x*,*y*)) respectively, the coefficient of MI, *MI(I,J)* can be mathematically expressed by:

$$MI(I,J) = \sum_{a,b} p_{IJ}(a,b) \ln \frac{p_{IJ}(a,b)}{p_I(a)p_J(b)} \cdot \quad (3)$$

where $p_{IJ}(a,b)$ corresponds to the joint probability distribution of *I* and *J* and $p_I(a)$ and $p_J(b)$ are the marginal probabilities distributions of gray values *α* of image *I* and *b* of image *J*, respectively. A disadvantage of MI is that it usually presents many local extremes in which the optimization procedure may be trapped, which reduces registration efficiency and reliability. Furthermore MI based registration schemes are sensitive to the used interpolation method and their accuracy is limited due to the discrete nature of *MI(I,J)* [1].

Chamfer Distance (CD): Two binary contour images are precisely aligned when the mean chamfer distance between them is minimum. 2D chamfer distance (CD) is computed by applying a suitable mask [6]. Usually the referenced contour distance map is computed prior to registration and used as a look-up table during the optimization procedure, in order to reduce execution time. Mean CD minimization is independent of the images gray level variances. However registration based on distance map is efficient when the image that contains the most contour information is assumed as the reference image [5].

Optimization Strategies

Downhill Simplex method (DSM), due to Nelder and Mead [7], is mostly recommended on applications that require execution speed. In this work, DSM was implemented as presented in [7]. The termination criterion was set equal to 10^{-6}, while the search space was restricted to [-0.1,+0.1] for scaling, [-6°,+6°] for rotation and [-150,+150] pixels for translation parameters around the initial specified position $\mathbf{P}_{o=}(a_1,a_2,a_3,a_4,d_x,d_y)=(1,0,0,1,0,0)$ in each of the parameters directions separately.

Powell's direction set method (PM) finds the minimum of the similarity function in the *N*-dimensional parameter space, by iteratively minimizing the function in one direction along the set of *N* conjugate different directions. However PM may be trapped to a local and not the global minimum of the function. In the present work, PM was implemented as described in [7]. The initial set of directions was considered to be the basis vector in each dimension and the parameters were optimized in the order $(d_y,d_x,\alpha_4,\alpha_3,\alpha_2,\alpha_1)$. The search space and the termination criterion were determined as in DSM implementation.

Simulated Annealing method (SA) is commonly used in registration applications to extract similarity function's global minimum hidden among many local minima [5]. It has been successfully applied in retinal images in combination with correlation and mutual information [1][3]. The concept of the method relies on thermodynamics' laws and depends on the mathematical expression of the similarity function, the generation function of the random steps, the acceptance criterion and the annealing schedule [7]. In this work, the random steps were generated from a uniform function and were added to the function value. The used annealing schedule was defined by $T=T_o/1.25$, where $T_o=0.1$ and $k_{max}=100$, where T_o is the initial temperature and k the number of iterations. Because of its stochastic nature, SA algorithm was followed by PM, which provides more stable outputs.

3 Experimental Results

In this work the different registration schemes were assessed on 23 retinal image pairs, 18 temporal RF pairs and 5 FA-RF pairs. The temporal RF images were taken up to five years apart. The four common optimization techniques, DSM, PM, DSM-PM and SA-PM were combined with each of the three similarity measures; C_{cc}, MI and CD. PM was implemented after DSM or SA. Every registration algorithm was initialized so that a_1, $a_4 \in$ [-1.01,1.01], a_2, $a_3 \in$ [-0.1,0.1], and dx, $dy \in$ [-150,150] pixels. Bilinear interpolation was used and checkerboard images of the registration were produced to allow visual assessment of every method. As reference image was taken the one that had more edge information, including noise.

The mean value and the variance of every similarity function for every optimization method were calculated. For each similarity function, a pairwise comparison of the optimization methods was performed by means of Student's paired t-test. The null hypothesis was that the optimization methods under comparison did not differ as per the value of similarity function. Results are shown in Tables 1-3.

The registration schemes were also compared to the gold-standard (manual) registration, which was performed by an expert. From every image pair, six pairs of bifurcation points were chosen, according to which the affine transformation parameters were calculated using the Least-Squares method (LSM). This procedure was repeated three times. The best set of parameters was chosen as the one that corresponded to the smallest associated Root Mean Square Error (RMSE) value. The average RMSE of the LSM for all image pairs was 0.77 pixels, a rather low value that shows the good accuracy during pair points definition.

For the evaluation of the similarity measures, one thousand edge points of the image-to-be-transformed from each pair were randomly chosen. The mean Euclidean Distance (RMSE) between the manual and automatic registered points was computed. In Fig.1 the mean RMSE and in Table 4 the mean, medium and maximum RMSE [8] of C_{cc}, MI and CD for DSM, PM, SA-PM, DSM-PM are presented. Finally a pairwise comparison of the registration errors was performed by means of Student's paired t-test. The null hypothesis was that the similarity measures under comparison did not differ as per the value of RMS errors, namely present the same registration accuracy. Results are shown in Table 5.

Table 1. C_{cc} mean value and variance, p-value, execution time and mean number of iterations for DSM, PM, SA-PM, DSM-PM for the 18 temporal image pairs

	Mean	Variance	p-value				Time (s)	Iterations
			DSM	PM	SA-PM	DSM-PM		
DSM	0.7775	0.1387	-	0.4918	0.4989	0.4993	29.9	396
PM	0.7785	0.1392	0.4918	-	0.4989	0.4996	119.9	1620
SA-PM	0.7786	0.1391	0.4907	0.4989	-	0.4993	92.2	1166
DSM-PM	0.7785	0.1391	0.4913	0.4996	0.4993	-	79.9	962

Table 2. MI mean value and variance, p-value, execution time and mean number of iterations for DSM, PM, SA-PM, DSM-PM for the 23 image pairs.

	Mean	Variance	p-value				Time (s)	Iterations
			DSM	PM	SA-PM	DSM-PM		
DSM	0.7887	0.3184	-	0.4589	0.4623	0.4571	45.3	354
PM	0.7549	0.3431	0.4589	-	0.4966	0.4982	186.0	1434
SA-PM	0.7976	0.3177	0.4623	0.4966	-	0.4948	171.8	1542
DSM-PM	0.7989	0.3161	0.4571	0.4982	0.4948	-	163.7	1289

Table 3. CD mean value and variance in pixels, p-value, execution time and mean number of iterations for DSM, PM, SA-PM, DSM-PM, for the 23 image pairs

	Mean	Variance	p-value				Time (s)	Iterations
			DSM	PM	SA-PM	DSM-PM		
DSM	8.0141	3.7033	-	0.4920	0.4832	0.4978	12.6	517
PM	7.9919	3.7270	0.4920	-	0.4754	0.4942	24.7	1580
SA-PM	8.0614	3.8782	0.4832	0.4754	-	0.4810	19.8	1336
DSM-PM	8.0079	3.7094	0.4978	0.4942	0.4810	-	14.1	1217

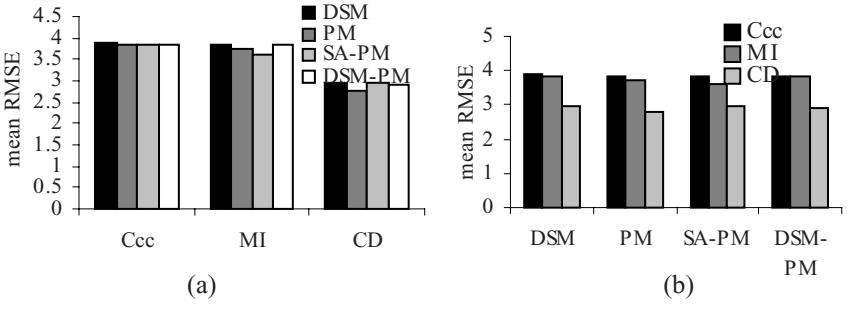


Fig. 1. (a) Mean RMSE of C_{cc}, MI and CD for DSM, PM, SA-PM, DSM-PM. (b) Mean RMSE of DSM, PM, SA-PM, DSM-PM for C_{cc}, MI and CD

According to Tables 1-3, and Fig. 1(a) no significant differences were observed (p-value>0.05, for all cases). All methods present the same registration accuracy. DSM

is the fastest method and needs the lowest number of iterations. PM seems to be very slow because of the small steps it takes in the parameter space. DSM-PM presents an average performance of DSM and PM, as far as execution time and number of iterations are concerned. Almost all optimization techniques depend on the shape of the similarity measure. If it has many extremes, then the optimization algorithms must be initialized close to the best solution. PM showed the strongest dependence on the initial guess when it was combined with the MI, because of the deviation of this similarity function surface from quadratic form. SA-PM was almost independent from the initial guess, since it represents a global optimization technique.

Table 4. Mean , medium (med) and maximum (max) RMSE in pixels, for C_{cc}, MI, CD combined with DSM, PM, SA-PM, DSM-PM, for all image pairs

	C_{cc}			MI			CD		
	mean	med	max	mean	med	max	mean	med	max
DSM	3.897	1.777	19.57	3.837	2.857	17.52	2.968	1.823	11.45
PM	3.849	1.840	19.35	3.731	2.296	17.38	2.783	1.488	9.67
SA-PM	3.828	1.834	19.56	3.602	2.162	17.54	2.969	2.252	10.91
DSM-PM	3.864	1.777	19.57	3.829	2.291	17.52	2.899	1.823	9.87

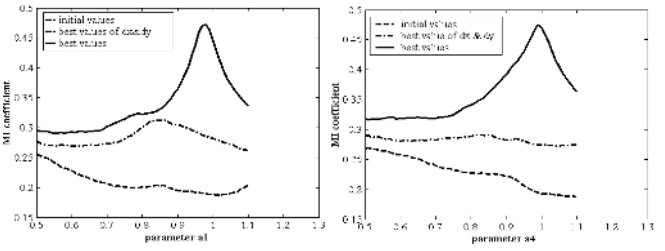
Table 5. P-value for C_{cc}, MI, CD for all registration cases. The numbers in bold correspond to accepted p-value for the null hypothesis.

	p-value								
	mean			med			max		
	Ccc	MI	CD	Ccc	MI	CD	Ccc	MI	CD
Ccc	-	**0.07**	$3x10^{-5}$	-	**0.06**	$4x10^{-4}$	-	$2x10^{-7}$	10^{-4}
MI	**0.07**	-	10^{-5}	**0.06**	-	0.008	$2x10^{-7}$	-	$2x10^{-4}$
CD	$3x10^{-5}$	10^{-5}	-	$4x10^{-4}$	0.008	-	10^{-4}	$2x10^{-4}$	-

C_{cc}, seems to be very efficient combined with PM because it has a unique extreme. However C_{cc} did not succeed in registering FA-RF image pairs, due to the nonlinear dependence between the gray levels of the two images. Also C_{cc}, according to Fig.1 (b) and Tables 4 and 5 presents high registration errors in comparison with CD, due to small but existent deviation of the images gray levels dependence from linearity, probably because of noise.

MI coefficient presents many local extremes and has different shape with different values of the transformation parameters. An example is presented in Fig.3, which shows the dependence of MI on the affine model's parameters a_1 and a_4 for one of the examined image pairs, when the other parameters were kept constant. Only near the best solution presents MI a global extreme. According to Table 5 MI presents almost the same registration accuracy with C_{cc}. MI coefficient is well coupled with SA-PM, which does not present a strong dependence on the initial guess.

CD does not depend on the gray levels of the images. It seems to be a robust and fast similarity measure that, according to Fig.1 can be combined well with all optimi-

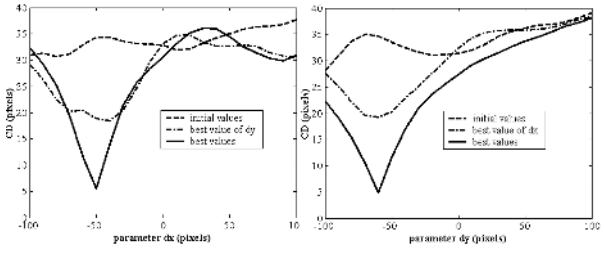


(a) (b)

Fig. 2. Dependence of MI coefficient on (a) a_4 and (b) a_1 transformation parameters, for a temporal RF image pair registration. Solid curves correspond to the best values of a_2, a_3, d_x, d_y and a_4 or a_1, dashdot curves to best values only of d_x and d_y, while in dashed curves the stable transformation parameters values were far from the best.

zation techniques, when contour extraction from images is possible. In Table 3 the required time includes the segmentation time interval as well. CD is the most accurate similarity measure. However CD depends strongly on translation parameters initial values. An example is shown in Fig. 4, that presents CD dependence on d_x and d_y parameters for another image pair.

For all the 23 examined cases, the error of the manual method was far and away smaller than those of the automatic techniques (mean RMSE=0.77, med RMSE=0.83 and max RMSE=1.14 pixels). Because of this and the fact that the placement of external markers in retinal image registration applications is not possible due to the great sensitivity of the human eye, manual registration was considered as a gold-standard procedure. In Table 4 the obtained errors of the automatic techniques include the error of the manual technique as well.

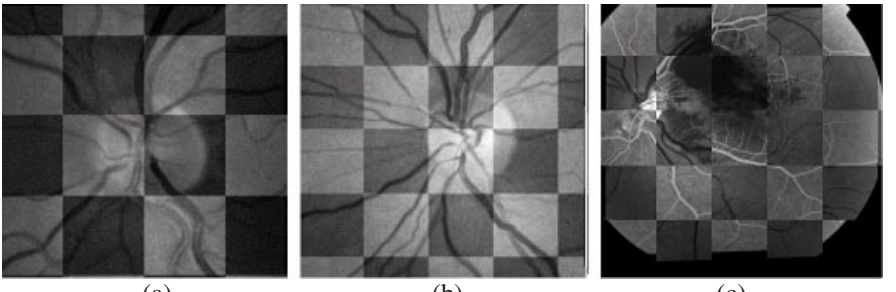


(a) (b)

Fig. 3. Dependence of CD on (a) *dy* and (b) *dx* transformation parameters, for a temporal RF image pair registartion. Solid curves correspond to the best values of a_1,a_2, a_3, a_4,and *dx* or *dy*, dashdot curves to best values only of *dx* or *dy*, while in dashed curves the stable transformation parameters values was far from the best.

Finally the affine model seems to be adequate for registering retinal images that are misaligned because of changes in the position between the camera and the patient. Deformable models could offer further improvements in the registration final result, especially in the case of glaucoma, where visual evaluation of vessels deformations in the area of the optic disk is essential.

Fig. 4 shows three different retinal image pairs, randomly chosen from the 23 pairs, registered with different schemes. As it can be seen from the images there is absolute continuity between vessels, something that shows the success of the registration schemes.



(a) (b) (c)

Fig. 4. (a) Chessboard image of a temporal RF image pair registered by PM combined with C_{cc}. (b) chessboard image of a temporal RF image pair registered by SA-PM coupled with MI. (c) chessboard image of an FA-RF image pair registered with DSM-PM combined with CD

4 Conclusion

In this work different global and local automatic registration schemes were applied to temporal RF and to FA-RF retina image pairs. The different techniques were compared on and evaluated against manual registration. The comparison showed that chamfer distance is an accurate and cost effective similarity measure that can be combined with common optimization techniques.

References

1. N. Ritter, et al "Registration of Stereo and Temporal Images of the Retina", IEEE Trans. on Medical Imaging, vol. 18, No. 5, May 1999.
2. F.Laliberte, L. Gagnon, "Registration and Fusion of Retinal Images-An Evaluation Study", IEEE Trans. on Medical Imaging, vol. 22, No. 5, May 2003.
3. G. Matsopoulos, N.A. Mouravliansky, K.K. Delibasis and K.S. Nikita, "Automatic Retinal Image Registration Scheme Using Global Optimization Techniques", IEEE Trans. on Inf. Tech. in Biomedicine, vol 3, No 1, March 1999.
4. L. Shaphiro, G Stockman: Computer Vision, Prentice Hall, New Jersey, 2001.
5. J. Mainz and M. Viergever, "A survey of medical image registration", Medical Image Analysis, vol.2, No 1, 1998.
6. G. Borgefors, "Distance Transformations in Arbitrary Dimensions", Com. Vision Graphics and Image Processing, vol. 27, 1984.
7. "Numerical Recipes in C: The Art of Scientific Computing", 1992 Cambridge University Press.
8. J. West, et al "Comparison and Evaluation of Retrospective Intermodality Brain Image Registration Techniques", J Comput Assist Tomogr, vol. 21, No 4, 1997.