14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

IMPROVING QUALITY OF AUTOFLUORESCENCE IMAGES USING NON-RIGID IMAGE REGISTRATION

Kubecka L., Jan J., Kolar R., Jirik R.

Department of Biomedical Engineering, Brno University of Technology
Brno University of Technology, 612 00, Brno, Czech Republic
phone: + (420) 541143603, fax: + (420) 541149542, email: libor@kubecka.info
web: www.dbme.feec.vutbr.cz

ABSTRACT

This work concerns quality improvement of auto-fluorescence retinal images by averaging of non-rigidly registered images. The necessity of using the elastic spatial transformation model is documented as well as the need for similarity criterion capable of dealing with the non-homogenous and variable illumination of retinal images. The presented multilevel registration algorithm provides parameters of primarily affine and then B-spline free-form spatial transformation optimal with respect to the mutual information similarity criterion. The registration was tested on three modeled image sets of 100 images. The difference of artificially introduced pre-deformation displacement field and the displacement field found by our algorithm clearly showed the ability to compensate for the diverse modeled distortions. Further, the registration algorithm was used for improving quality of realistic retinal images using averaging of registered frames of image sequences. The whole method was verified by processing of 16 time series of real images. The gain in signal to noise ratio in the averaged registered images with respect to individual frame reach the expected about 4dB, without introducing a visible blur. The final image was substantially less blurred than the non-registered averaged image, which is documented by comparison of the autocorrelation functions of both images.

1. INTRODUCTION

In the recent decades, correlation between the distribution of retinal dye lipofuscin (LF) in the retinal pigment epithelium (RPE) and various ophthalmic diseases was proven in several studies. According to [1], accumulation of lipofuscin in the RPE is associated with degeneration of RPE cells and photo-receptors and therefore it has a strong diagnostic value. In [2] the relation of the size of LF areas to glaucoma development was shown.

Distribution of LF can be measured *in vivo* by excitation of RPE using 488nm blue argon laser and subsequent detection of auto-fluorescence radiation of LF making use of confocal scanning laser ophthalmoscope (Heildelberg Retina Angiograph, HRA) with a front-mounted broad band pass filter with a lower wavelength cutoff of 500 nm excluding reflected laser light.

Unfortunately, since the auto-fluorescence is of low intensity, the pure HRA scans are of rather poor signal to noise ratio (SNR_{single_frame}=22.3dB according to definition (8)) and can not be used for reliable determination of the LF distribution directly. This situation is caused particularly by comparable value of intensity of AF radiation and of thermal noise of the used semiconductor detector. Our possibilities to improve the situation were limited to posterior data processing, namely averaging of time sequences of images of the identical scene. We used time series of 9 pure HRA scans provided by the system. With respect to the eye motion, some type of flexible registration had to be used before averaging. This paper deals with the non-rigid multi resolution registration method that proved capable of preserving the original spatial resolution in averaged images with correspondingly improved SNR.

2. METHODS

Image registration can be treated as an optimization problem with the goal of finding the spatial mapping that will bring the moving image into alignment with the fixed image. It can be formalized as finding the optimal parameter vector $\boldsymbol{\alpha}_0$ of spatial transform $T_{\boldsymbol{\alpha}}$,

$$\boldsymbol{\alpha}_0 = \arg\min_{\boldsymbol{\alpha}} C\left(f(\mathbf{x}), m\left(T_{\boldsymbol{\alpha}}(\mathbf{x})\right)\right), \qquad (1)$$

where f is the fixed (reference) image and m is the floating image to be registered, which is transformed by $T_{\boldsymbol{\alpha}}$ to coordinates of the fixed image. The registration quality, corresponding to the transform $T_{\boldsymbol{\alpha}}$ spatially transforming moving image using given parameters $\boldsymbol{\alpha}$ is evaluated by the global similarity criterion C. The optimal transformation $T_{\boldsymbol{\alpha}0}$ transforms the moving image m into the image $m(T_{\boldsymbol{\alpha}0}(\mathbf{x}))$, which is maximally similar to the fixed image f, thus minimizing C.

2.1 Similarity metric

Even if the registration problem is mono-modal, the non-homogenous illumination differing among the images prevents us from choosing a simple similarity criterion such as the sum of squared differences or the normalized cross-correlation. Substantially better results were reached with the mutual information (MI) similarity criterion originally designed for multi-modal problems. MI is able to overcome non-homogenous illumination, which produces unequal intensity of corresponding pixels. Let $p_f(f(\mathbf{x}))$ a $p_m(m(\mathrm{T}(\mathbf{x})))$ are the probabilities of occurrences of brightness values in

the images and let $p_{fm}(f(\mathbf{x}),m(\mathrm{T}(\mathbf{x})))$ be the joint probability, then the mutual information can be defined as follows:

$$MI(f,m)\Big|_T = \\ = \sum_{\mathbf{x}} p_{fm}[f(\mathbf{x}),m(T(\mathbf{x}))]\ \log\frac{p_{fm}[f(\mathbf{x}),m(T(\mathbf{x}))]}{p_f[f(\mathbf{x})]\,p_m[m(T(\mathbf{x}))]} \tag{2}$$

The choice of MI as the similarity criterion was supported by availability of an effective implementation of the criterion using sc. Parzen windowing method [7], which gives us access to analytical directional derivatives of the criterion. These are necessary for efficient optimization of large scale problems like non-rigid registration.

2.2 Optimizers

Three optimizers in combination with the multi-resolution approach were used in order to find the global extreme of the MI similarity criterion:

- **Controlled random search** (CRS) is a kind of contraction process where an initial sample set of points is iteratively contracted by replacing the worst point with a better one. See [8], [9] for a greater detail.
- **Powell's method** computes step directions for function minimization as conjugate directions. See [10] for a greater detail.
- **Limited memory Broyden, Fletcher, Goldfarb, Shanno method** (L-BFGS) is a quasi-Newton optimization algorithm. It uses function values and gradients to build up a picture of the surface to be optimized. See [5] for a greater detail.

2.3 Spatial transformation

The distortions introduced during the acquisition of HRA time sequence are produced either by global eye motion between subsequent scans, or because of motion during scanning of an individual HRA image. In order to compensate for the distortion, we separated our model of geometric distortion into global and local motion parts so that it can be formalized as

$$T(x,y) = T_{global}(x,y) + T_{local}(x,y). \tag{3}$$

The global component is expressed by

$$T_{global}(x,y,1) = \begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 & \alpha_5 \\ \alpha_6 & \alpha_7 & \alpha_8 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{4}$$

where the world coordinates of the transformed point are (x'/w, y'/w). For the case of affine global transform the parameters $\alpha_6 = \alpha_7 = 0$ and $\alpha_8 = 1$.

The local motion is modeled using B-spline based free-form deformation model [4]:

$$T_{local}(x,y) = \sum_{l=0}^{3}\sum_{m=0}^{3} B_l(u)\,B_l(v)\,\mathbf{\Phi}_{i+l,\,j+m}, \tag{5}$$

where

$$i = \lfloor x/n_x \rfloor - 1,\ j = \lfloor y/n_y \rfloor - 1, \\ u = x/n_x - \lfloor x/n_x \rfloor, v = y/n_y - \lfloor y/n_y \rfloor \tag{6}$$

n_x, n_y being the numbers of control (grid) points, $\mathbf{\Phi}$ are parameters of the transform defining spatial shift of control points ($\mathbf{\Phi}$ is denoted as $\boldsymbol{\alpha}$ for the optimization), B_l is a cubic B-spline basis function.

$$B_0 = (1-u^3)/6 \\ B_1 = (3u^3 - 6u^2 + 4)/6 \\ B_2 = (-3u^3 + 3u^2 + 3u + 1)/6 \\ B_3 = (u^3)/6 \tag{7}$$

2.4 Optimization strategy

We used multi-resolution pyramidal approach for the optimization and different types of optimizers at each level according to dimensionality of the optimization problem and to the risk of getting stuck in the local optimum. In the first step (rough optimization), optimal parameters of translation transform are found first using images sub-sampled into the resolution 256x256 pixels. Then searching for affine transform parameters is done making use of results of the previous step. In these two steps, controlled random search algorithm (CRS) is used as a robust optimization method for this optimization of a low-number of parameters. Further, we assume that found parameters are close enough to the global optimal parameters so that the similarity metric can be treated as approximately quadratic form and therefore the Powell method of conjugate directions can be used in order to quickly find the assumed global optimum. In all these steps, standard mutual information metric evaluated from mutual histogram is used. The mutual histogram is directly computed from the nearest-neighbor interpolated images (while using CRS) or from 2nd-order B-spline interpolated images (at Powell).

Further (fine optimization), optimal shifts of control points of the B-spline spatial transform are sought. In this phase, we used L-BFGS optimization algorithm capable of efficiently optimizing high number of parameters. Multi-resolution was applied in image domain (down-sampling images) and in spatial transformation domain (various densities of grid control points). Thus, optimization is firstly done using images with resolution 512x512 with the spatial transformation based on 8x8 control points (128 parameters to optimize) and then with spatial transformation based on 26x26 control point (1352 parameters to optimize). The dimensionality of the parametric space is large but the used B-spline spatial transform has an advantage of the local property: one control point affects only its nearest neighborhood [3] and consequently, the parameters are partially separable, which enables effective optimization.enclosed by square brackets (e.g., [2]).

3. EXPERIMENT

3.1 Image Data

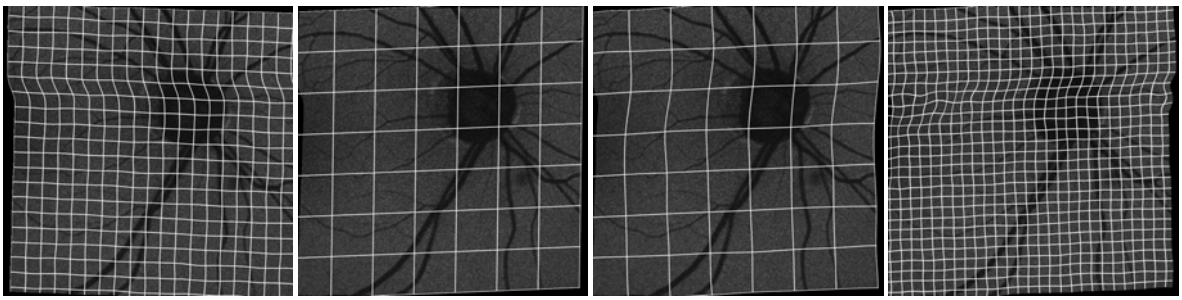
Images were acquired using a confocal scanning laser ophthalmoscope (Heildelberg Retina Angiograph, HRA) in Fluorescein Angiography (FA) mode. Series of nine images were taken for each of 16 patients. Images are of 20°x20° field of view and have the spatial resolution 10 μm/pixel. Images were acquired in two resolutions: 1024x1024 pixels, 5 μm/pixel and 512x512, 10 μm/pixel [6].

Bound	Transl. [px]		Scale [%]		Shear	
	x	y	x	y	x	y
Upper	35	35	0.7	0.7	0.05	0.05
Lower	-35	-35	-0.7	-0.7	-0.05	-0.05

Table 1: Range of affine transform parameters used for creating pre-deformed image sets.

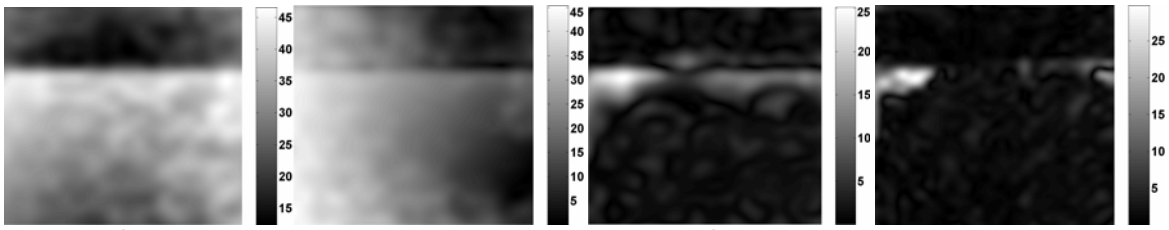
	SET I	SET II	SET III
Mean value of the pre-displacement field norm (DF) [px]	32.8585	34.1771	35.0628
Mean value of the affine registered DF norm [px]	33.1668	29.5343	34.9714
Mean value of the coarse registered DF norm [px]	4.6898	10.6015	2.9809
Mean value of the fine registered DF norm [px]	3.7141	9.0282	2.6642

Table 2. Mean values of the mean value of error deformation field (DF) norm across the whole testing set. DF is a vector field describing spatial shift of each image point. Error DF is computed as the sum of the pre-deformation field and the post-registration field.



A B C D

Figure 1: A - Pre-deformed image (20x20 control points) from the SET III. B - affine registered image. C - coarsely non-rigidly registered image (8x8 control points). D - finely non-rigidly registered image (28x28 control points).



A B C D

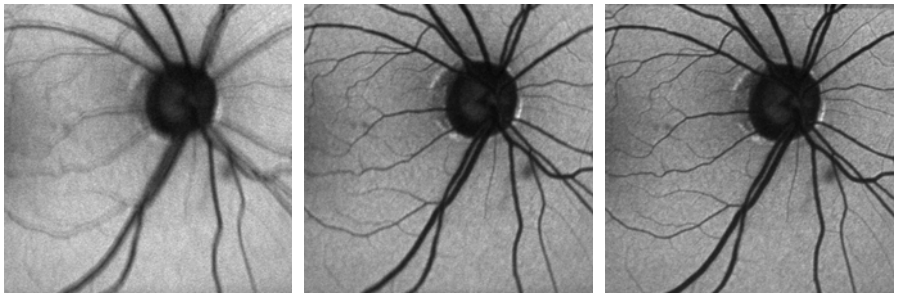
Figure 2: A – Norm of the displacement field (DFN) of pre-deformation transform (image of Fig. 1, from the SET III); Mean value of the norm is 30.235 pixels. B – Norm of the error displacement field (EDFN) after the affine registration; mean(EDFN) = 21.3858px. C – EDFN after the coarse non-rigid registration (8x8 control points); mean(EDFN) = 3.206px; D – EDFN after the fine non-rigid registration (28x28 control points); mean(EDFN) = 2.3866px. See text for the detailed description.

Scan time is about 100ms for one high resolution scan and 50ms for low resolution.

Individual images are highly corrupted by noise and by variable non-homogenous illumination; therefore, there is a need for preprocessing images before registration.

3.2 Testing

We chose a randomly selected HRA high-resolution image of the size 1024x1024 pixels and derived a new image set containing 100 artificially deformed images. For this purpose, parameters of the affine transform were selected randomly from the constrained parameters space (see Tab.1). The local distortion was modeled by shifting nodes of the underlying grid containing 20x20 control points. Two image sets with different range of the shift were made. In the SET I, the maximal range of shift was 15 points in both axes. While for the SET II we used the range of 35 pixels. Since this model of image distortions is not physically based, a spatial model of acquisition was used for creating SET III; here, one part of the image was randomly shifted with respect to the other part and small random jitter of the control nodes in 10 pixels range was added (see Fig. 1).). Then, the proposed registration algorithm was performed and a displacement vector field (DF) containing the spatial shift of each image pixel was calculated from the parameters of the spatial transform found by each registration level. The quality of the final registration was measured by comparing these DF and the known DF of pre-deformation. After registration, the found DF should be inverse to the introduced DF when it is ideally compensated. Therefore the error DF was computed as the sum of the pre-deformation DF and the registration DF which should be ideally zero. The norm of this error field was calculated as a measure of the displacement remaining after compensating the pre-deformation by the computed registering transform. The mean value of this norm across the whole testing set for each registration level is given in the table 2. In the figure 2, examples of magnitude of pre- and post registration displacement vector field for one image from the SET III are given.



A B C

Figure 3: A - Average image without registration. B - Average image with rigid registration is still slightly blurred. C - Average image with non-rigid registration.

As can be seen from the Fig 2A, the introduced spatial distortion has a two clearly separated parts resulting from the random shift of one part of the image with respect to the other. Fig. 2B shows that affine transform is not able to substantially compensate the distortion where the distortion on right part of the image is suppressed but it is enhanced on the left part. The ability of the proposed non-rigid algorithm to remove the distortion is visible in Fig 2C,D where displacement errors after two levels of elastic registration are depicted. We can see that the displacement error field is almost zero except to some location where insufficient information was contained in the image so that the MI criterion was not able to evaluate the similarity reliably.

4. RESULTS

The proposed algorithm was used for aligning 15 time series, containing 9 images each. Seven time series were in the resolution of 1024x1024 pixels, others was 512x512 pixels. In all cases the proposed algorithm was successful. The quality of registration was estimated subjectively by constructing movies each containing all images from every series and detecting possible misalignment as motion. Another test of quality was done when comparing the quality of the averaged images before and after alignment (see Fig. 3).

The average signal to noise ratio (SNR) of a single image in the sequence was computed and compared to the SNR of the image combined without registration, after rigid registration and after elastic registration (see Table 3). An assumingly constant background area Ω was selected using image thresholding and morphological operations, and SNR was computed as the ratio between the global signal range (maximum minus minimum image) and standard deviation over Ω,

$$SNR = 20\log_{10}\left(\operatorname{range}\left\{ f(\mathbf{x}) \right\} \Big/ \operatorname{std}\left\{ f(\mathbf{x})\Big|_{\Omega} \right\} \right). \quad (8)$$

All images were normalized to 256 grey-levels prior to SNR computation hence the range was constant and the SNR was dependent on noise variation only. The average gain in the SNR was around 4dB, which compares well with the theoretical optimum 4,77dB for nine totally independent images.

The sharpness of the resulting images was assessed by the main lobe width of the 2D autocorrelation function; the wider the lobe the less sharp the image. The criterion is the count of discrete autocorrelation-function values higher than 3/4 of the function maximum; thus lower value means sharper image.

Image	SNR [dB]	Sharpness [pts]
Single frame	23.3454	123.2
No registration	27.1868	5095.8
Rigid	28.3543	2361.4
Elastic	27.8664	1679.9

Table 3: Average signal-to-noise ratio and sharpness measure for the image series.

The results are summarized in the Table 3. Seemingly, the single frame (non-combined) image is the sharpest by far; however, the result is influenced by a high level of wide-band noise narrowing substantially the autocorrelation main lobe. Thus any comparison is only possible among combined images with a similar SNR; comparing those three, it is clearly seen that the best result by far is obtained by combining flexibly registered images of the sequences.

5. CONCLUSIONS

A method for improving quality (SNR) of auto fluorescent retinal images by their elastic alignment is presented. The method is based on maximization of mutual information (MI) and consists of low- and high-resolution steps. The low resolution step compensates for the global mis-alignment caused by patient's inter-image movements. In order to avoid local extremes of MI, the stochastic controlled random search optimization routine is used for finding the optimal affine transformation model on this level.

The high-resolution step eliminates the local distortions caused by eye movements mainly during single image acquisition using B-spline based free form deformation model and efficient quasi-Newton L-FBGS optimizer.

The algorithm was widely tested on the image sets containing 100 artificially deformed images. These tests proved the necessity for using elastic registration in order to be capable of compensating distortions present in the processed retinal images. Further, the algorithm was used for alignment of 15 time-series each containing 9 real patient images. All of these images were successfully registered. The quality of the regis-

tration was evaluated in the averaged images which were less noisy than the single frames by an average gain in signal to noise ratio about 4dB, as expected. The substantially lower blurring of the averaged images after elastic registration compared to averaged images without registration or even after rigid registration is documented.

6. ACKNOWLEDGEMENT

Authors sincerely acknowledge the contribution of Dr. R. Laemmer and Dr. Ch. Y. Mardin, Augen-Klinik Erlangen (Germany) who provided the data from HRA and valuable consultations. This project was supported by the research centre DAR (Ministry of Education, Czech Republic), proj. no. 1M6798555601 and also by the grant no. CZE MS 0021630513 of the Ministry of Education, Czech republic.

REFERENCES

[1] C.K. Dorey, G. Wu, D. Ebenstein, et al. "Cell loss in the aging retina. Relationship to lipofuscin accumulation and macular degeneration", *Invest Ophthalmol Vis Scie*, vol. 30, pp. 1691–1699, 1989.

[2] A. Viestenz, C.Y. Mardin, A. Langenbucher, G.O. Naumann. "In-vivo Measurement of Autofluorescence of Parapapillary Atrophic Zone of Optic Discs with and without Glaucomatous Athrophy", *Klin Monats Augenheil*, vol. 220, no. 8, pp. 545-50, Aug 2003.

[3] J. Kybic, M. Unser, "Fast Parametric Elastic Image Registration", *IEEE Trans Image Processing*, vol. 12, no. 11, pp. 1427-1442, Nov 2003.

[4] D. Rueckert. L. I. Sonoda, C. Hayes, D.L. Hill, M. O. Leach, D. J. Hawkes. "Nonrigid registration using free-form deformations: application to breast MR images", *IEEE Trans Medical Imaging*, vol. 18, no. 8, pp. 712-21, Aug 1999.

[5] D. C. Liu, J. Nocedal, "On the limited memory BFGS method for large scale optimization", *Mathematical Programming*, vol., 45, pp. 503-528, 1989.

[6] HRA 2 Brochure, available on http://www.heidelbergengineering.com

[7] P. Thevenaz, M. Unser, "Optimizatiobn of Mutual Information for Multiresolution Image Registration", *IEEE Trans Image Processing*, vol. 9, no. 12, pp 2083-2099, Dec 2000

[8] J. Tvrdik: "Generalized controlled random search and competing heuristics", *In MENDEL 2004, 10th Internetional Conference on Soft Computing*. pp. 228-233, University of Technology, Brno, 2004.

[9] M. M. Ali, A. Torn, S. Viitanen: "A Numerical Comparison of Some Modified Controlled Random Search Algorithms", *J. Global Optimization*, no. 11, pp. 377-385, 1997

[10] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, second edition, 1992.