Elastic Registration for Auto-fluorescence Image Averaging

Libor Kubecka, Jiri Jan, Radim Kolar, Radovan Jirik

Abstract — **The paper describes restitution of geometrical distortions and improvement of signal-to-noise ratio of auto-fluorescence retinal images, finally aimed at segmentation and area estimation of the lipofuscin spots as one of the features to be included in glaucoma diagnosis. The main problems – geometrical and illumination incompatibility of frames in the image sequence and a non-negligible "shear" distortion in the individual frames – have been solved by the presented registration procedure. The concept and some details of the MI-based regularized registration, together with evaluation of test results form the core of the contribution.**

I. Introduction

In recent studies, a correlation between the distribution of retinal dye lipofuscin (LF) in the retinal pigment epithelium (RPE) and various ophthalmic diseases has been shown, as e.g. in Dorey et. all [1] who found a strong diagnostic value - specially for glaucoma - of accumulation of lipofuscin in the RPE associated with degeneration of RPE cells and photoreceptors. Distribution of LF on retina can be measured in-vivo using confocal scanning laser ophthalmoscope in a special mode. Unfortunately, images produced this way have a rather poor signal to noise ratio. To improve it, we applied averaging to time sequences of (almost) identical retinal images. Due to patient eye movements during measurement and also due to scanning method of acquisition, there are often misalignments throughout the image sequence, as well as geometrical "shear" distortions in individual images that have to be compensated for prior to averaging. This paper deals first with the non-rigid registration method we developed reflecting the particularities of this concrete application. Secondly, the results of the spatial registration of sequential images (including image rectification) will be shown and discussed.

II. Data Acquisition

Retinal images were acquired using a confocal scanning laser ophthalmoscope (Heildelberg Retina Angiograph, HRA) in Fluorescein Angiography (FA) mode. Images of 20° x 20° field of view were acquired in two resolutions, either 1024 x 1024 pixels, 5 μm/pixel or 512 x 512 pixels, 10 μm/pixel. Scan time is about 100ms for a high resolution scan and 50ms for a lower resolution one. A nine-image sequence was taken for each of 16 patients.

Due to acquisition conditions in clinical practice, the obtained images are strongly corrupted by noise and non-homogenous illumination, besides to motion artefacts.

Manuscript received April 3, 2006. This project is running under the support of the research centre DAR (sponsored by the Ministry of Education, Czech Republic), project no. 1M6798555601.

L.Kubecka, J.Jan, R.Kolar and R.Jirik are with the Department of Biomedical Engineering, FEEC, Brno University of Technology, Brno, 61200, Czech Republic (corresponding author L.Kubecka, phone: +420 54114-3603; fax: +420 54114 9542; e-mail: kubecka@feec.vutbr.cz)

III. HRA Image Registration

Image registration throughout the sequence is treated as an optimization problem [3], aiming at spatial mappings aligning the to-be-registered (floating) images with the chosen fixed reference image. For a particular image couple, it can be formalized as

$$\hat{\mathbf{\Phi}} = \arg\left\{\min_{\mathbf{\Phi}} C\left(f_R(\mathbf{x}), f_T\left(\mathbf{g}(\mathbf{x};\mathbf{\Phi})\right)\right)\right\}, \qquad (1)$$

where $f_R(\mathbf{x})$ is the reference image and $f_T(\mathbf{x})$ is the image to be registered via the mapping $\boldsymbol{g}(\mathbf{x};\mathbf{\Phi})$. The optimal transformation defined by $\hat{\mathbf{\Phi}}$ transforms the image $f_T(\mathbf{x})$ into the image $f_T\left(\mathbf{g}\left(\mathbf{x};\hat{\mathbf{\Phi}}\right)\right)$ best fitting $f_R(\mathbf{x})$ thus minimizing the criterion C. The registration quality, corresponding to the current vector $\mathbf{\Phi}$ of transform parameters, is evaluated by the optimization criterion $C(..)$, the construction of which turned out crucial for the successful registration of HRA sequences.

A. Optimization criterion

The optimizing registration criterion, as designed for this task, consists of two parts: a similarity metrics, evaluating the match of both images, and a regularization term preventing artefactual behaviour of the registration process.

Similarity metrics: Experiments have shown that – in spite of monomodality of the problem - simple metrics like Euclidean distance or normalized cross-correlation fail due to non-homogenous illumination, which produces incomparable intensities of corresponding pixels, and also due to poor signal to noise ratio (SNR) in the registered images. Therefore, the mutual information (MI), evaluating general non-linear dependence of two data sets and thus usually used for multi-modal similarity evaluation, turned out the only suitable choice even for the mono-modal HRA.

MI is the measure of statistic relation between the fixed and moving image, based on entropies expressed by the joint probabilities $p(\dots,\dots)$ of pixel intensities:

$$S(f_R, f_T, \mathbf{\Phi}) = -\sum_{\mathbf{x}_i}\left\{ p\left[f_R(\mathbf{x}_i), f_T\left(\mathbf{g}(\mathbf{x};\mathbf{\Phi})\right)\middle|\mathbf{\Phi}\right] \log_2 \frac{p\left[f_R(\mathbf{x}_i), f_T\left(\mathbf{g}(\mathbf{x};\mathbf{\Phi})\right)\middle|\mathbf{\Phi}\right]}{p_R\left[f_R(\mathbf{x}_i)\right]\cdot p_T\left[f_T\left(\mathbf{g}(\mathbf{x};\mathbf{\Phi})\right)\middle|\mathbf{\Phi}\right]}\right\}. \qquad (3)$$

Following Mattes [4], we used Parzen windowing method for the joint and marginal probabilities estimation. Here, the probability density function (PDF) is construed from intensity samples with superimposed kernel function w. Let L_R and L_T be the discrete sets of intensities associated to the reference and floating images and let $w(\xi)$ be a separable Parzen window satisfying $w(\xi) > 0,\ \forall \xi$; then we define the joint discrete Parzen probability of simultaneous occurrence of intensities κ in reference image and ι in floating image transformed by geometrical transform with parameters $\mathbf{\Phi}$ as

$$p(\iota,\kappa;\mathbf{\Phi}) = \alpha(\mathbf{\Phi}) \sum_{\mathbf{x}_i \in \Omega} w\left(\kappa - f_R(\mathbf{x}_i)\right) w\left(\iota - f_T\left(\mathbf{g}(\mathbf{x}_i,\mathbf{\Phi})\right)\right), \quad (4)$$

where Ω is the image domain and $\alpha(\mathbf{\Phi})$ is a normalization term, which ensures that

$$\sum p(\iota,\kappa) = 1, \text{ for } \iota \in L_T, \kappa \in L_R. \quad (5)$$

The marginal discrete probabilities are then computed by summing up the joint probabilities [4]. The B-spline kernel function has been selected for the Parzen window w; it has the advantage in having a finite support resulting in faster computation, and further it gives the direct access to the metrics derivatives, which is needed in the optimization.

Regularisation term: Experiments have shown that HRA image registration based on pure MI similarity criterion lead often to a not well-physically-founded spatial transformation, especially in noisy regions on higher resolution levels of flexible registration. To prevent such behaviour, we introduced two regularization terms into the criterion. First, the incompressibility constraint has been set up, similarly like in [5], which prevents changing of area of local imaged structures. In a small neighbourhood of the point **x**, the local compression or expansion produced by the transform **g** is give by the determinant of the transform Jacobian:

$$\mathbf{J_g}(\mathbf{x}) = \det\left(\begin{bmatrix} \partial g_1(\mathbf{x})/\partial x_1 & \partial g_1(\mathbf{x})/\partial x_1 \\ \partial g_2(\mathbf{x})/\partial x_1 & \partial g_2(\mathbf{x})/\partial x_2 \end{bmatrix}\right). \quad (6)$$

The value of $\mathbf{J_g} > 1$ ($\mathbf{J_g} < 1$) signalizes local expansion (compression) at the point **x**, while $\mathbf{J_g} = 1$ for incompressible transform. Therefore, we define the energy term penalizing the distance of $\mathbf{J_g}$ from 1 for every image point **x**

$$E_{\text{area}} = \frac{1}{\text{count}(\Omega)} \sum_{\mathbf{x}\in\Omega} \left|\log\left(\mathbf{J_g}(\mathbf{x})\right)\right|, \text{ where } \text{count}(\Omega) = \sum_{\mathbf{x}\in\Omega} 1, \quad (7)$$

Ω being the image domain.

The second introduced constraint evaluates smoothness or roughness of the transformation, generally as

$$E_{rough} = w_d \int_\Omega \left\|\nabla \operatorname{div} \mathbf{g}\right\|^2 dx_1 dx_2 + w_r \int_\Omega \left\|\nabla \operatorname{rot} \mathbf{g}\right\|^2 dx_1 dx_2. \quad (8)$$

We have used a version of this constraint simplified by neglecting the second derivatives of combined g_1, g_2, thus

$$\begin{aligned} E_{\text{smooth}} = &\int_\Omega \left[\left(\frac{\partial^2 g_1(\mathbf{x})}{\partial x_1^2}\right)^2 + 2\left(\frac{\partial^2 g_1(\mathbf{x})}{\partial x_1 \partial x_2}\right)^2 + \left(\frac{\partial^2 g_1(\mathbf{x})}{\partial x_2^2}\right)^2 \right] dx_1 dx_2 \\ &+ \int_\Omega \left[\left(\frac{\partial^2 g_2(\mathbf{x})}{\partial x_1^2}\right)^2 + 2\left(\frac{\partial^2 g_2(\mathbf{x})}{\partial x_1 \partial x_2}\right)^2 + \left(\frac{\partial^2 g_2(\mathbf{x})}{\partial x_2^2}\right)^2 \right] dx_1 dx_2 \end{aligned} \quad (9)$$

This constraint simulates stress of a stretched elastic material, similarly as in the thin-plate splines.

The MI metrics and both regularization constraints have been combined into the final optimization criterion

$$C_{similarity}(f_R, f_T, \mathbf{\Phi}) = S(f_R, f_T, \mathbf{\Phi}) + w_v E_{\text{area}} + w_s E_{\text{smooth}}, \quad (10)$$

where w_v and w_s are weights of the area or smoothness penalization, respectively. In order to allow compensation of small area changes caused e.g. by vessels pulsation, the area constraint must not be too strong; experimentally, w_v=0.0001 has been found suitable. As the B-spline transformation model (see later) ensures smoothness of the deformation on coarse resolution levels of registration, w_s is non-zero for the high-resolution transformation only, where it is set to w_s=0.01.

B. Models used in registration process

Spatial transform model: The transformation performs spatial mapping of pixels from the reference (fixed) image $f_R(x_1,x_2,t_0)$ taken in the time t_0 into the coordinate system of the floating image $f_T(x_1',x_2',t)$ taken in time t of the dynamic image sequence. This transformation model follows the principle of backward brightness mapping.

Distortions during the acquisition of HRA image sequence can be caused either by global eye motion between subsequent scans or also in frame of a single image due to a movement of the patient's eye within a single HRA scan. Then, a part of the scan is shifted with respect to other part(s). Consequently, we have separated our model of geometric distortion into global and local motion components as follows:

$$\mathbf{g}(x_1,x_2) = \mathbf{g}_{global}(x_1,x_2) + \mathbf{g}_{local}(x_1,x_2). \quad (11)$$

The global part is supposed to consist of affine transformation, which has 6 degrees of freedom in 2D, namely spatial shifting, skewing and scaling. In generalized coordinates, the affine transform of a point (x,y) is defined as

$$g_{global}(x_1,x_2,1) = \begin{bmatrix} x_1' \\ x_2' \\ w \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}, \quad (12)$$

where the world coordinates of the transformed point are $(x_1'/w, x_2'/w)$. For affine global transform $a_{31}=a_{32}=0$, $a_{33}=1$.

The local part of the transformation model is a free form deformation (FFD) based on B-spline basis functions, widely used for the purpose of image registration, e.g. [5], [6]. These bases have finite support hence they can be controlled locally, which results in lower interdependency of transformation parameters and more efficient optimization. We define the local B-spline transformation as

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} + \sum_{\mathbf{k}\in\mathbb{Z}^N} \mathbf{\Phi}_{\mathbf{k}} \beta^{(n)}\left(\frac{\mathbf{x}}{\mathbf{h}} - \mathbf{k}\right) = \left[g_1(\mathbf{x}), g_1(\mathbf{x}), \ldots\right], \quad (13)$$

$$g_i\left(\left[x_1, x_2\right]\right) = \sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} \Phi_{k_1,k_2} \beta^{(m)}\left(\frac{x_1}{h_1} - k_1\right) \beta^{(m)}\left(\frac{x_2}{h_2} - k_2\right), \quad (14)$$

where **h** is the knot spacing, *m* B-spline order and **k** index vector of a control point. **Φ** denotes the $n_1 \times n_2$ grid of control points $\mathbf{\Phi}_{\mathbf{k}}$ with uniform spacing **h.** These parameters control the transform and can be interpreted as spatial shifts of the control points.

Brightness interpolation model: We assume discrete image $f(\mathbf{x}_i)$ defined on a set of points of regular grid $\mathbf{x}_i$, then the interpolation term providing the continuous representation of the image necessary for spatial image transformation is $f(\mathbf{x})$ and can be described as follows:

$$f(\mathbf{x}) = \sum_i c(\mathbf{x}_i) \beta^{(n)}(\mathbf{x} - \mathbf{x}_i), \qquad f(\mathbf{x}_i) = f(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}_i}, \quad (2)$$

where, in the 2D case, $\beta^{(n)}(\mathbf{x}) = \beta^{(n)}(x_1) \cdot \beta^{(n)}(x_2)$ is a separable n^{th} order B-spline convolution kernel and $c(\mathbf{x}_i)$ are their respective weighting coefficients. This B-spline interpolation scheme allows us to compute image derivatives required for efficient computation of criterion derivatives in optimization.

C. Optimization strategy

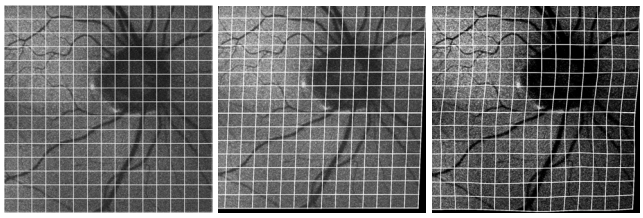
We have used a multilevel (multi-resolution) approach to optimization, with different types of optimizers on individual levels - according to dimensionality of the optimization problem on the level, and also with respect to the risk of getting stuck in a local minimum.

First, images are sub-sampled to the resolution 128x128 pixels and then rough translation parameters are found. As the next step, the affine transform parameters are searched for, using the results of the previous step as initiation. In these two steps, the *controlled random search* (CRS) optimizing algorithm is used, being a contraction process where an initial set of *N* points in **Φ** space is iteratively contracted by replacing the worst point with a better one; for details see [3]. In both steps, the nearest-neighbour (0-order B-spline) interpolation of transformed image values has been found (surprisingly) as best for the criterion computation.

After the two steps, the found parameter vector may be assumed as close enough to the global optimum so that the similarity metrics can be considered a quadratic form and consequently, the Powell's method may be used to quickly improve precision of the globally optimal **Φ** of affine transformation. This method finds the *N*-dimensional minimum of *C* by repeatedly minimizing *C* in a single dimension, gradually along *N* different directions; it only requires evaluations of *C* itself and not of its gradient (see [7] for details). In this step, 2nd-order B-spline brightness interpolation is used in the criterion computation. No regularization (w_a=w_r=0) was used in the affine registration step.

As the next step, flexible registration is applied using images down-sampled to 256x256 pixels. First, a coarse registration with spatial transform controlled by 11x11 control-nodes is done. Here we make use of the limited-memory Broyden, Fletcher, Goldfarb and Shannon method (L-BFGS); a quasi-Newton unconstrained nonlinear optimizer, utilising both function values and gradients to build up a picture of the surface to be optimized (see [8] for details). 2nd order B-splines for image value interpolation and regularised optimum criterion with w_s=0.0001, w_v=0 were used.

Finally, the found transform has been up-sampled to FFD controlled by 31x31 points, initiating the last step of elastic registration, which is done with this B-spline resolution. It requires 1922 parameters of the FFD to be optimized with respect to the criterion *C* using the regularization weights set to w_s=0.01, w_v=0.0001. The typical results of registration process are shown on the Fig. 1.



A B C

Figure 1: A - Control grid before registration. B - Control grid after the affine registration. C - Control grid after flexible registration showing restitution of a shear distortion of a frame.

TABLE 1: DESIGN OF THE FINAL MULTI-RESOLUTION ALGORITHM

Level - Transform	Sub-Sampl.	Optimizer	Interpol	w_v	w_s
1-translation	4 (8)	CRS	0	0	0
2-affine	4 (8)	CRS	0	0	0
3-affine	2 (4)	Powell	2	0	0
4-spline 11x11 nodes	2 (4)	L-FBGS	2	0	10^{-4}
5-spline 31x31 nodes	2 (4)	L-FBGS	2	10^{-4}	10^{-2}

IV. AVERAGING OF HRA SEQUENCES

The proposed registration algorithm has been used for aligning image sequences each containing 9 images acquired from 15 patients. Of this data, 7 sequences were of 1024x1024 pixel resolution, others were of 512x512 resolution. In all cases the registration using the proposed algorithm has been successful. The quality of the registration was evaluated both subjectively via movies of each sequence after registration, detecting any misalignment as a movement, and objectively. The quantitative evaluation of the registration quality has been done on the registered images combined into an image by one of the three techniques: averaging, the principle component analysis (PCA) with taking the first component [9], and the minimum noise fraction transform (MNF) [9] also with the first component. The average 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. An assumingly constant background area Ω was selected using image thresholding and morphological operations, and SNR was computed as the ratio between signal range (maximum minus minimum image value) and standard deviation over Ω,

$$SNR = 20\log_{10}\left(\frac{\text{range}\{f(\mathbf{x})\}}{\text{std}\{f(\mathbf{x})\}}\right). \quad (16)$$

The average gain in the SNR was around 4dB, which compares well with the theoretical optimum 4,77 dB for nine totally independent images. All tested combination methods have given similar results: SNR_{aver}=26.9393, SNR_{PCA}= 26.9383, SNR_{MNF}= 26.9448. Therefore, the simplest of them, i.e. averaging was used for the final algorithm. The resulting images in different phases of processing are on the Fig. 2.

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.

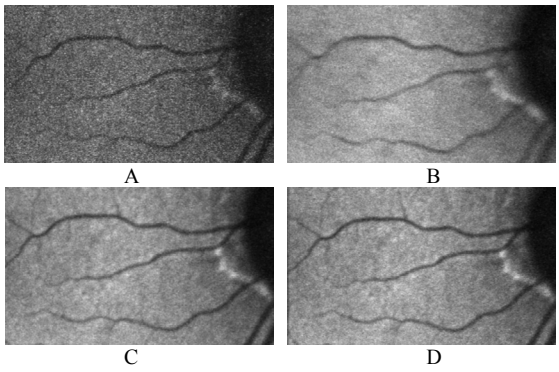


Fig. 2: A - sub-image of a single frame, B - averaged sub-image without registration, C - averaged sub-image with rigid registration (still slightly blurred). D - and after fkexible registration (see narrow vessels)

TABLE 2: AVERAGE SIGNAL-TO-NOISE RATIO AND SHARPNESS MEASURE FOR THE IMAGE SERIES.

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

The results are summarized in the Table 2. Seemingly, the single frame (non-combined) image is 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.

V. DISCUSSION AND CONCLUSIONS

The method of registration and averaging has met the goal – providing low-noise images with a good resolution and corrected movement artefacts, as it has been shown quantitatively in the previous two sections. The results could have only be achieved using relatively advanced means, in spite of the seeming simplicity of the task.

The presented method of flexible alignment of retinal auto-fluorescence images is based on maximization of mutual information (MI) because the involved noise and especially inhomogeneous and inconsistent illumination do not allow using standard mono-modality similarity measures. The levelled approach uses different resolutions and also different optimisation strategies at the individual levels; this allows for the optimal computational speed and also prevents trapping of the solution in a local extreme. On higher levels of computation, the MI similarity optimization criterion is complemented by two weighted regularization terms.

The algorithm has been successfully tested on 15 image sequences each containing 9 images. In each of these sequences, its images were after registration combined into a low-noise, low distortion image. The quality of registration has been quantitatively evaluated via a sharpness criterion – the main lobe width of the 2D correlation function.

ACKNOWLEDGMENT

Authors sincerely acknowledge the contribution of Dr. R. Laemmer and Dr. Ch. Y. Mardin, Augen-Klinik, University Hospital Erlangen (Germany) who provided HRA data and valuable consultations.

REFERENCES

[1] 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 Monatsbl Augenheilkd*, vol. 220, no. 8, pp. 545-550, August 2003.

[2] P. Perona, J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE *Trans. on Pattern Anal. Mach. Intell.*, vol. 12, pp. 629–639, 1990.

[3] L. Kubecka, J. Jan, "Registration of Bimodal Retinal Images – improving modifications," in *Proc. 26th Internat. Conf. IEEE-EMBS*, San Francisco, USA, 2004, pp. 1695-1698.

[4] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "Non-rigid multimodality image registration," In *Medical Imaging 2001: Image Processing*, pp 1609–1620, 2001.

[5] T. Rohlfing, C.R. Maurer, D.A. Bluemke, M.A. Jascobs, "Volume-Preserving Nonrigid Registration of MR Breast Images Using Free-Form Deformation With an Incompressibility Constraint," *IEEE Trans. On Medical Imaging*, vol. 22, no. 6, pp. 730-741, June 2003.

[6] Kybic, J, Unser, M. Fast Parametric Elastic Image Registration. IEEE Transactions on Image Processing, vol. 12, no. 11, November 2003.

[7] W. H. Press, S. A. Teukolsky, W.T. Vetterling, B. P. Flannery *Numerical Recipes in C*, Cambridge University Press, 1995.

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

[9] K.B. Hilger, "Exploratory Analysis of Multivariete Data," PhD. Thesis, TU of Denmark, Nov. 2001.