Retinal image registration using creases as anatomical landmarks

David Lloret, Joan Serrat, Antonio M. López, Andrés Soler, Juan J. Villanueva
Computer Vision Center and Departament d'Informàtica,
Universitat Autònoma de Barcelona,
Edifici O, 08193–Cerdanyola, Spain.
Tel. +34 935811863, Fax +34 935811670
E–mail : david@cvc.uab.es

Abstract

Retinal images are routinely used in ophthalmology to study the optical nerve head and the retina. To assess objectively the evolution of an illness, images taken at different times must be registered. Most methods appeared so far have been designed specifically for a single image modality, like temporal series or stereo pairs of angiographies, fluorescein angiographies or SLO images, which makes them prone to fail when conditions vary. In contrast, the method we propose has shown to be accurate and reliable on all *the former modalities. Actually, it has been adapted from the 3D registration of CT and MR image to 2D. Relevant features (also known as landmarks) are extracted by means of a robust creaseness operator, and resulting images are iteratively transformed until a maximum in their correlation is achieved. Our method has succeed in more than 100 pairs tried so far, in all cases including also the scaling as a parameter to be optimized.*

1. Introduction

Retinal images are a diagnostic tool for examining the condition of the retina. For a number of diseases, e.g. diabetic retinopathy, it is convenient to track the evolution through a period of time, while for others it may be necessary to compare images taken at the same time but with different modalities. These images show the vascular tree of the eye, which permits to determine objectively the areas where blood flow seems occluded or to be leaking. There are several retinal image modalities: **retinographies** are taken using the ophthalmoloscope under natural light, and often they have been applied a green filter (**green images**) to discard the dominant red component. **Fluorescein angiographies** sense the fluorescence emitted from the vessels of the retina after the injection of a contrast dye. Another novel technique is the Scanning Laser Ophthalmoscope **SLO**, which provides better resolution, controllable excitation intensity and variable depth of focus. All this modalities can be combined to produce simultaneous **stereo pairs** or sequential **temporal series** of the same eye, in order to study the optic disk for signs of progression of an illness. Martínez-Costa [6], for instance, seeks high increments in gray level at the foveal center in a series of temporal images to detect macular leakage due to retinal vein occlusions.

Since the position of the eye changes every time a picture is taken, the resulting images will show some degree of rotation and translation. Before comparing them, one has to align their corresponding features. This problem, known as registration, has been widely studied in medical imaging. See, for instance [5] for a survey of this techniques.

A number of papers are devoted solely to detect the vessels in the image. In [9] Tolias et al. employ fuzzy clustering to determine vessel and non-vessel regions, while in [7] Martínez-Pérez binds a region growing algorithm to local curvature. The information thus extracted is useful for clinical analysis, but it is not used for comparison to other images. In the registration field, another group of papers extract the bifurcation points of the vessels by means of mathematical morphology [11] or Förster's detector [1] and then they choose some heuristic invariant to rotation and translation to match corresponding points. Thus, the resulting method seems to be highly dependent on the results of the extraction step. Whenever the number of detected points is too low, or the corresponding sets too separate, their algorithm is prone to fail.

Another well known approach in medical image registration are the mutual information methods. These methods, which were applied firstly on the registration of head images in 1995 [10], [4], do not extract corresponding features but make use of all the information available. They measure the statistical dependence or information redundancy between the image intensities of the corresponding voxels of both images. Nicola Ritter has report success in a method based

0-7695-0750-6/00 $10.00 ● 2000 IEEE

on this approach in [8]. In her paper, she provides with an excellent review of existing methodologies, and she stresses the importance of a good database of images to test the algorithms under realistic assumptions. Indeed, some papers take into account only translations, or accept as a good result a ten-pixels misregistration. A major part of the paper is dedicated on the problem of local maxima. They occur when the function to measure the alignment gets stuck into a value which is the highest in a local neighbourhood but not in the while parameter space. Because the optimization is a function of 3 parameters (two translation plus one rotation angle, plus sometimes two scaling factors) an exhaustive search is not feasible. Thus she uses a scheme based on pyramidal sampling and the simulated annealing search algorithm. She reports success on her experiments on an about 100 images database.

Our major criticism to her work regards to the proposed search method. Although successful, the simulated annealing algorithm is highly dependent on a large number of parameters, 6 global plus 7 for each layer of the pyramidal search. These parameters are hidden deep in the code, making it impossible for a user to tune to any change in the image constraints. Also, because the method is not designed to segment the contents of the images, it can not provide with any further measure about the vascular tree.

2. Our method

Our method resembles the human approach to image matching in the sense that we also employ as guidelines features common to both images. It seems natural to us not to restrict the comparison to the bifurcation points, but to use the more significant structures visible in the image, that is, the vessel tree. Opposite to the bifurcation points approach ([1], [11]), ours will not have such a strong dependence of the quality of the segmentation.

Perfect segmentation of the whole vessel tree is not an easy task because images often provide poor contrast and vessels vary in diameter and intensity level; some papers are dedicated to this sole purpose ([7], [9]). However, we were fortunate to have already addressed a similar problem for CT and MR brain registration. For this purpose, we developed a precise and reliable detector of the creaseness (ridgeness or valleyness) of an image. We describe this operator in section 2.1.

After vessels have been extracted, we needed to choose a scheme to transform iteratively the images until they become aligned. In literature many methods to measure the alignment exist; we could, for instance, segment each branch and follow some heuristic to choose its corresponding. Its drawback is the same as bifurcation methods have: the extra segmentation step is prone to propagate its errors to the optimization step. Therefore, we choose the well-known cross-correlation between the two images. This measure is fast and also, as we require, it matches common extracted features despite of missing or spurious data.

We have implemented the iterative Simplex algorithm to optimize the alignment process, and added an initial wide search to improve its robustness. The optimization process is fully described in section 2.2.

2.1. Creaseness measures

Vessels are reliable landmarks in retinal images because they are almost rigid structures and they are depicted in all modalities. Moreover, they can be thought as creases (ridges or valleys) when images are seen as landscapes. Amongst the many definitions of crease, the one based on level set extrinsic curvature (LSEC) has useful invariance properties. Given a function $L : \mathbf{R}^d \rightarrow \mathbf{R}$, the level set for a constant l consists of the set of points $\{\mathbf{x}|L(\mathbf{x}) = l\}$. For $2D$ images, L can be thought as a topographic relief or landscape and the level sets are its level curves. It is well–known that negative minima of the level curve curvature κ, level by level, form valley curves, and positive maxima ridge curves.

However, the usual discretization of LSEC is ill defined for a number for cases, giving rise to discontinuities in the output image. Instead, we have employed the $MLSEC - -ST$ operator, as defined in [3] ($\tilde{\kappa}_4$ in page 130).

2.2. Search scheme

After the creaseness extraction, the following step is to iteratively transform one of the images until it becomes properly aligned with the other. A suitable function to measure the quality of the alignment is the correlation function $C_T = \sum_{\vec{x} \in f} f(\vec{x}) \cdot g(T(\vec{x}))$, where f and g are the creaseness images and T represents a transformation whose parameters we want to test. A key step is not to transform all the pixels in the image, but only those with creaseness values higher than a small fixed threshold. This step saves up to 95% of the total computations.

The function C_T together with the 3 parameters of the transformation defines a space where the maximum value, i.e. the best alignment, must be found. The Simplex algorithm works properly for this purpose, but for a number of images it fails to converge. The problem is that the initial guess for these images had been set too far to the final value, and the Simplex algorithm gets trapped somewhere in a local maximum. Therefore, we need to scan fast and efficiently the search space to provide the Simplex convergence algorithm with a set of candidate seeds, one of which might be sufficiently close to the global maximum. Since we do not need high accuracy for this initial step, we have used the same cross correlation function but in the Fourier domain. First, we transform one of the images for a given

rotation angle and we keep the best of the candidate translations given by the Fourier cross correlation. After repeating this step for a range of rotations, we will have a set of candidate values to be optimized.

For large images (more than 512 rows or columns) the proposed method has a bottleneck at the initial step: each sampled rotation demands the computation of three costly Fourier transforms. The solution was to sample the images in a hierarchical scheme down to a size of 128 × 128 following a scheme we applied to head images registration in a previous paper [2], this is, building two pyramids where the creaseness images are at the bottom, and the intensity of each pixel at a given level is the maximum of a local neighbourhood at the previous level. Compared to the standard Gaussian sampling, this approach has the advantage to keep the creases through the hierarchy.

The explained scheme worked successfully for all sets of images; even without the initial seed step, all except one actually converged. Despite of the global convergence, results were visually disappointing when branches were located far from the center. We found this was caused by a different scaling in the images, and we decided to include scaling in a final additional step. Results then were accurate for all parts of the image.

3. Results

We designed a test-bench to validate the choices we took while designing the search strategy. We evaluated five search configurations: **(1)** no hierarchical, no initialization **(2)** no hierarchical, initialization using best seeds of a single Fourier search **(3)** same as previous, but hierarchic and doing Fourier at the coarse level **(4)** hierarchic, search for 10 rotations at 5 deg each **(5)** same as previous, 26 rotations at 2 deg. The structure of the experiment was as follows: first, we register a given pair of images. Then, we misalign one images by known parameters T_T, and the registration algorithm is applied again. To measure the distance from the given to the recovered parameters T_T, we have taken the mean distance after applying the two corresponding transformations to all the non-void pixels of the creaseness image, this is, those detecting vessels. We applied 100 different transformations to four pairs of images, taken randomly within the range of translations of $\pm 25\%$ the size of the image in pixels, rotations of ± 25 deg and scaling $\pm 10\%$. Table 1 shows the results on a PC *Pentium* at 350 MHz with 256 MB of memory running under Linux.

It is clear that a hierarchical approach is necessary to achieve some robustness, and that seeds propagate well the results of the search from the broad to the fine hierarchical levels. In addition to robustness, this approach reduces in one order of magnitude the search time. Please note that trial transformations are actually far more demanding than real cases: normal registrations have very little scaling, and translations and rotation angles are lower. When we performed tests which did not include scaling to the same images, results were almost 100% successful. For all tests, the recovered transformations had a mean error lower than $0,2$ pixels, which shows that our alignment measure is properly defined.

	Image 1 370 × 278		Image 2 250 × 202		Image 3 460 × 416	
	Rec (%)	Time	Rec (%)	Time	Rec (%)	Time
1	6	4,6	16	1,7	9	6,1
2	69	9,3	44	3,6	37	10,3
3	94	61,4	98	44,3	93	66,2
4	94	9,7	80	4,0	92	10,3
5	97	10,8	83	5,0	93	11,2

Table 1. Results of the registration algorithm for 100 random trial transformations for 3 pairs of images: (1) Angiography-retinography (2) Angiography-retinography with partial occlusion (3) Stereo angiographies. The Rec column represents the percentage of transformations successfully recovered (mean error lower than 5 pixels); the time is in seconds. Search methods as numbered are explained in section 3.

We have registered successfully more than 100 pairs of images of several modalities: 49 stereo and 48 temporal pairs (1 year apart, size 460 × 416) of angiographies and 10 green images to angiographies. Also, the same technique was used to register a temporal series of 140 SLO image which had been taken within 5 second, with a size of size 512 × 512. All the slices of this series were registered to the same last slice because it was the one which showed better the vascular tree. We did not keep information between registration runs. The method ran for all images unmodified for all but two parameters, which are related to the creaseness extraction: the scale of the crease and the smoothness of the image, and it had a typical response time of 10 seconds. The visual inspection of the results, which was easy to perform because creases overlap when properly registered (see figure 3), was fully satisfactory for all the sets. For the temporal series, despite the poor contrast of the initial slice, where little could be seen, the registration was not successful for only four pairs. As a future commercial application, we are setting the algorithm to reduce its computational charge and to decide automatically the success of a registration. Eventual sparse rejections are unimportant if we manage to make our application almost real-time.

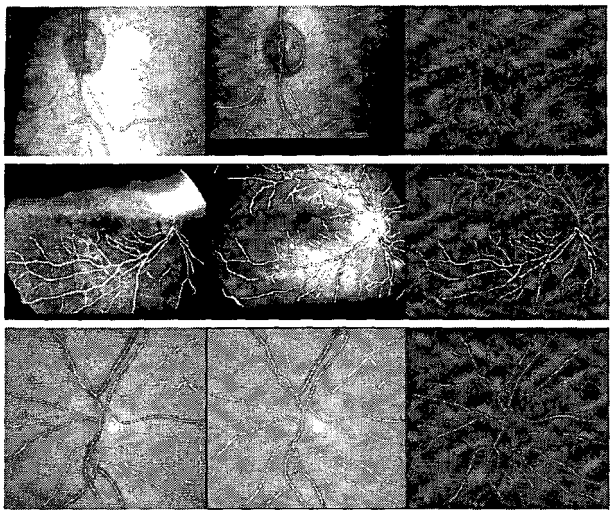

Figure 1. First and second columns: original images with creases superimposed in white. Third column: overlapping creases appear in white, non-overlapping in black and gray. From top to bottom: SLO–SLO, retinography to green, and retinography to retinography.

4. Conclusions

We have presented a novel method to register retinal images, with several advantages compared to others: a) reliable landmarks extraction using a creaseness operator b) cross-correlation for robustness against superfluous or missing landmarks c) hierarchical approach to speed up the results and improve robustness d) experiments include severe miss-registration conditions.

Our method works fast and reliably for the four retinal modalities. We have explored its employment for registering SLO temporal series, and first results have showed to be highly satisfactory. As a future work, we want to employ the vessels already extracted and registered to measure some clinically relevant indicators. Also, we are reducing the response time of the algorithm to make it suitable for clinical real-time applications.

5. Acknowledgments

This research has been partially funded by CICYT projects TIC97–1134–C02–02 and TAP96–0629–C04–03. We would like to thank the source of our images: F. Zana at the Centre de Morphologie Mathématique, Ecole des Mines de Paris, E. de Ves at Instituto de Robótica, Univ. de València and J. Pinós at Hospital General Universitario de Valencia. Special thanks to C. Barry and N. Ritter at the McCusker Glaucoma Foundation for letting us use their large image database. Finally, thanks to Manuel F. González. Penedo and M.J. Carreira at Universidade da Coruña for the SLO series.

References

[1] J. Domingo, G. Ayala, A. Simó, E. de Ves, L. Martínez-Costa, and P. Marco. Irregular motion recovery in fluorescein angiograms. *Patter recognition letters*, 18:805–821, 1997.

[2] D. Lloret, A. López, J. Serrat, and J. Villanueva. Creaseness-based CT and MR registration: comparison with the mutual information method. *Journal of Electronic Imaging*, 8(3):255–262, July 1999.

[3] A. López, D. Lloret, J. Serrat, and J. Villanueva. Multilocal creasness based on the level set extrinsic curvature. *Computer Vision and Image Understanding*, (77), 2000.

[4] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. on Medical Imaging*, 16(2):187–198, April 1997.

[5] J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[6] L. Martínez-Costa, P. Marco, G. Ayala, E. de Ves, J. Domingo, and A. Simó. Macular edema computer-aided evaluation in ocular vein occlusions. *Computers and biomedical research*, 31:374–384, 1998.

[7] M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton, and S. A. Thom. Retinal blood vessel segmentation by means of scale-space analysis and region growing. In C. Taylor and A. Colchester, editors, *Medical image computing and computer-assisted intervention-MICCAI'99*, volume 1679 of *Lecture notes in computer science*, pages 90–97. Springer, 1999.

[8] N. Ritter, R. Owens, and J. Cooper. Registration of stereo and temporal images of the retina. *IEEE Transactions on Medical Imaging*, 18(5):404–418, May 1999.

[9] Y. A. Tolias and S. M. Panas. A fuzzy vessel tracking algorithm for retinal images based on funzzy clustering. *IEEE Transactions on Medical Imaging*, 17(2):263–273, April 1998.

[10] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.

[11] F. Zana and J. Klein. A multimodal registration algorithm of eye fundus images using vessels detection and hough transform. *IEEE transactions on Medical Imaging*, 18(5):419–428, May 1999.