COVARIANCE-DRIVEN RETINAL IMAGE REGISTRATION INITIALIZED FROM SMALL SETS OF LANDMARK CORRESPONDENCES

C.-L. Tsai[1] *C. V. Stewart*[1] *B. Roysam*[2] *H.L. Tanenbaum*[3]

[1]Dept. of Computer Science
Rensselaer Poly. Inst.
Troy, New York 12180–3590
tsaic,stewart@cs.rpi.edu

[2]Dept. of ECSE
Rensselaer Poly. Inst.
Troy, New York 12180-3590
roysab@rpi.edu

[3]The Center for Sight
349 Northern Blvd.
Albany, NY 12204
Hltcfs2001@aol.com

ABSTRACT

An automatic retinal image registration algorithm would be an important tool for detecting visible changes in the retina caused by the progress of a disease or by the impact of a treatment. Developing such an algorithm is difficult, especially for feature-poor images of diseased eyes. In this paper, a new retinal image registration algorithm is described that bootstraps an estimate of the parameters of a high-order, inter-image transformation model based on just one or two initial retinal image landmark correspondences. Hypothesized sets of initial correspondences are obtained through invariant indexing. For each such set, an initial, low-order transformation covering a small image region is estimated. Sufficiently accurate initial estimates are gradually expanded to a high-order transformation that covers the entire retina using constraints generated by alignment of the vasculature. The expansion and switch in model orders is entirely driven by the covariance matrix of the estimated transformation parameters. The resulting algorithm registers images to accuracies of less than a pixel in just a few seconds.

1. INTRODUCTION

A fully-automatic algorithm for registering a pair of retinal images will have a substantial impact on analyzing, diagnosing and treating a number of diseases of the human retina (Figure 1). Potential applications include real-time tracking for treatment of blindness-causing conditions [1], change detection to measure the progress of a disease [2] or the impact of treatments [3], and multimodel integration to aid in diagnosis and surgical planning [4]. Solving the retinal image registration problem requires estimating the parameters of a transformation model mapping the coordinates of one image onto a second. This poses a number of significant challenges, several of which are illustrated in Figure 1: (1) The retina is a curved surface and usually no camera calibration parameters are available. (2) Illumination and camera settings such as focus and zoom can vary substantially between images. (3) Two images taken many months apparent can exhibit major physical changes in the structure and color of the retina. (4) Diseased eyes often have relatively few features on which to base registration. (5) Except under well-controlled conditions, the part of the retina visible in different images can differ substantially.

Several of these challenges are addressed in the literature. For example, most published techniques are feature-based — using the retinal vasculature, which is generally quite stable — to accommodate illumination variations between images [5, 6]. Our own recent work has introduced a quadratic inter-image transformation based on combining models of unknown retinal curvature, viewpoint changes, and camera parameters [7]. We have demonstrated a registration algorithm to estimate the parameters of this model using a hierarchical, robust technique. While this algorithm is fast and generally produces registration errors of less than one pixel on 1024×1024 images, it requires at least six landmark (branchings and cross-over points of the vasculature, illustrated in Fig. 2) correspondences. Images of a diseased eye often do not have nearly enough landmarks in common Fig 1. Thus, from a technical viewpoint, the fundamental challenge is "bootstrapping" reliable, accurate registration from as few landmark correspondences as possible.

This paper presents a new retinal image registration algorithm designed to address exactly this problem. The fundamental and novel ideas of the algorithm, illustrated in Fig. 3, are (1) to initialize the transformation estimate using a low-order model covering only a small region around a single landmark correspondence or a pair of landmark correspondences, and (2) to gradually extend and refine this to estimate a high-order, image-wide transformation. Crucially, the initial correspondences are insufficient to constrain the high-order transformation, but they are sufficient to initialize locally and at a low order. The processes of expanding the region covered by the transformation estimate and shifting to a higher-order model are both driven by analyzing the transformation uncertainty as measured by the

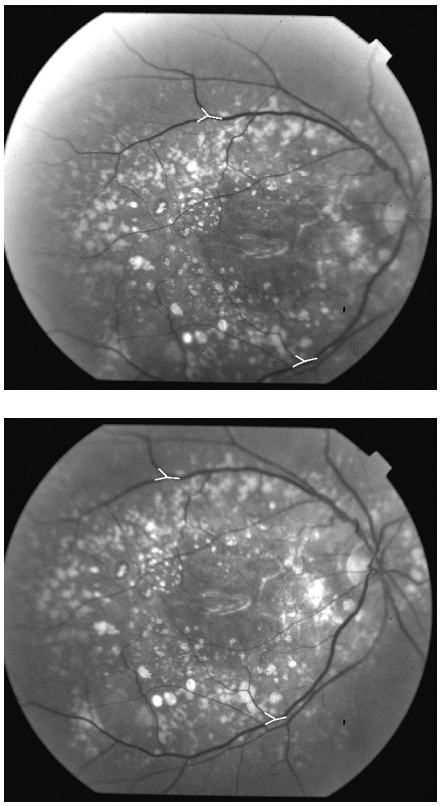

Fig. 1. Two different images of an unhealthy eye. The vasculature has poor definition and different parts of the eye are in focus in the two images. Automatic landmark extraction only produced two vascular landmark in common between the two images. These are drawn over top of the images.

covariance matrix of the estimate.

2. PRELIMINARIES

The new registration algorithm depends on a number of techniques and results published in our prior work. The landmarks used to establish initial correspondences are branching and cross-overs of the retinal vasculature. The constraints used in refinement are correspondences between centerlines of the retinal vasculature. Both are detected using a procedure that traces the retinal vasculature starting from seed points located along a series of vertical and horizontal grid lines [8].

The transformation describing the mapping of one image coordinate system onto the other is well-approximated by a 12-parameter, quadratic model [7]. Let $\mathbf{p} = (x, y)^T$ be the coordinates of a point location in one image, I_p, and let $\mathbf{q} = (u, v)^T$ be the coordinates of the same location on the retina in the second image, I_q. Then, the transformation

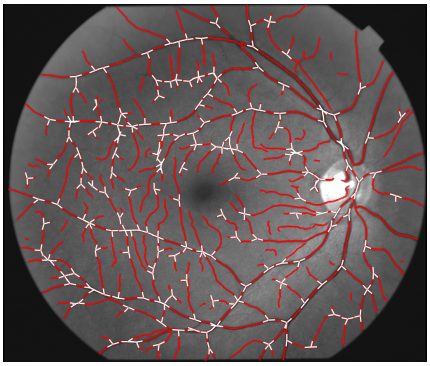

Fig. 2. Extracted vascular landmarks and vessel centerlines drawn on top of an image of a healthy eye. In images of diseased eyes, such as Fig. 3, these landmarks and centerline contours are much more sparsely distributed.

mapping $\mathbf{p}$ onto $\mathbf{q}$ may be written

$$\mathbf{q} = M(\mathbf{p}; \mathbf{\Theta}) = \mathbf{\Theta}(1, x, y, x^2, xy, y^2)^T \quad (1)$$

where $\mathbf{\Theta}$ is a 2×6 parameter matrix, which must be estimated. Importantly, this transformation is locally well-approximated by affine and even similarity transformations, which have 6 and 4 degrees of freedom, respectively, instead of 12.

3. GENERATING HYPOTHESIZED, INITIAL CORRESPONDENCES

The algorithm generates hypotheses for initial correspondences between a pair of landmarks or between two pairs of landmarks. The example in Figure 3 illustrates a single pair. Each hypothesis is used to estimate an initial similarity transformation, which is then refined and verified. If it is rejected, then a new hypothesis is generated. Hypothesis generation is the focus of this section.

Each detected vascular landmark is described using 8 features, as shown in Figure 4: the center location of the intersection region, three blood vessel orientations, and three blood vessel widths. Ratios between the widths and orientation differences do not change — are invariant — when a similarity transformation is applied to them. The orientations themseleves are invariant to translation and scaling. Thus, assuming small rotations, which is almost always true because of physical considerations, the three angles and two width ratios form an invariant signature. For pairs of landmarks that are sufficient near each other, a six-parameter similarity-invariant signature [9] is computed which does not involve vessel widths, the least accurate of the parameters. Thus, pair invariants allow for rotations and are less sensitive to feature instability.

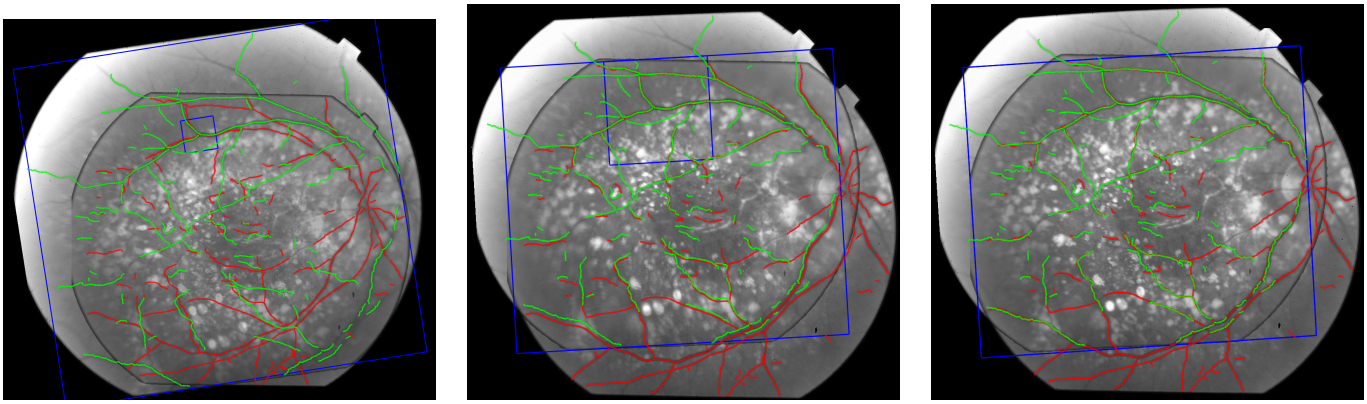

Fig. 3. An example illustrating the new registration algorithm. The left panel shows the initial alignment of the images from Fig. 1 using a similarity transformation computed from a single landmark correspondence. Extracted vascular centerlines from both images are shown (green and red). The vasculature is aligned in the small box around the landmark, but badly misaligned globally. Driven by the covariance matrix of the transformation estimate, the region is gradually expanded and the transformation refined (middle). When it is expanded to cover the entire overlapping region (right) the vasculature that is in common between the two images is accurately aligned.

The invariant signatures vectors are computed separately for each image. Hypotheses for correspondences are generated by finding signature vectors that are close to each other. Vectors are stored in a k-d tree for fast look-up. Hypotheses are ordered by chi-squared statistics on Mahalanobis distances between vectors.

Each hypothesized correspondence is used to compute a least-squares estimate of the similarity transformation aligning the surrounding image regions. The covariance matrix of the 4 similarity transformation parameters is computed as

$$\mathbf{\Sigma} = \sigma^2 \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}) \qquad (2)$$

where σ^2 is an error variance, and $\mathbf{H}^{-1}(\hat{\boldsymbol{\theta}})$ is the inverse Hessian matrix of the least-squares objective function evaluated at the vector of estimated similarity parameters, $\hat{\boldsymbol{\theta}}$.

This procedure leads to the initial estimate and region show in Figure 3(a).

4. COVARIANCE-DRIVEN REFINEMENT

The refinement procedure extends an initial transformation by increasing the region over which it is applied and increasing the model order until the transformation is clearly incorrect or until the region is image-wide and and the transformation accurately aligns the vasculature. This relies on the fact that an accurate alignment of the complex blood vessel pattern of the retina can not be accidental.

The procedure itself is most easily viewed as a generalization of an iterative-closest point algorithm [10]. Here is an outline of the steps, which start from an initial region $\mathbf{R}$,

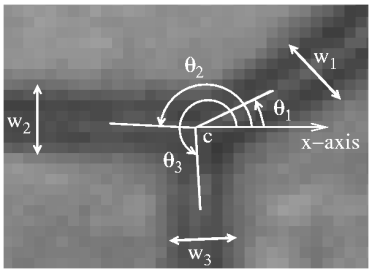

Fig. 4. Landmark features: $\mathbf{c}$ is the center location, θ_j are orientations of the three blood vessels that meet to form the landmark, and w_j are the vessel widths.

parameter estimate $\hat{\boldsymbol{\theta}}$, and covariance matrix $\mathbf{\Sigma}$. The mapping is from image I_p to I_q.

1. A sampling is taken of the vascular centerline (trace) points (Fig. 2) within region R of I_p. For each point $\mathbf{p}_i$ the current transformation estimate is applied to map $\mathbf{p}_i$ onto I_q, producing point $\mathbf{p}'_i$.

2. For each transformed point, $\mathbf{p}'_i$, the closest trace point, $\mathbf{q}_i$, on a vascular centerline contour in I_q is found.

3. The transformation parameter vector is re-estimation using robust estimation techniques [11] and constraints based on the match between the points $\mathbf{p}_i$ and $\mathbf{q}_i$. The constraints are the distances between the transformed $\mathbf{p}_i$ and a linearization of the contour at $\mathbf{q}_i$.

4. The new error variance is robustly estimated and the new covariance matrix is calculated as in (2).

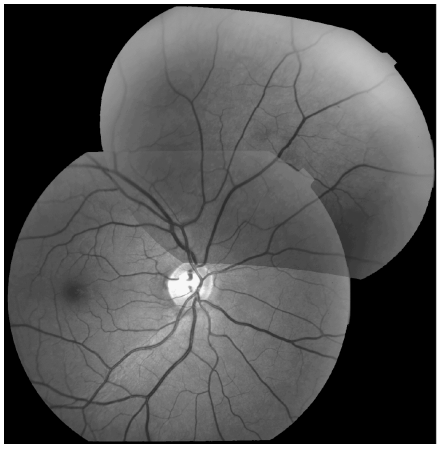

Fig. 5. Alignment of two images having only 25% overlap.

5. The region is expanded based on the "point transfer" [12, Ch. 4] of the transformation at the boundaries. For a point $\mathbf{p}$ on the boundary of R, let $\mathbf{J}_p$ be the Jacobian of the current transformation estimate with respect to $\mathbf{p}$. Then the transfer error at $\mathbf{p}$ is the point covariance matrix $\mathbf{\Sigma}_p = \mathbf{J}_p \mathbf{\Sigma} \mathbf{J}_p^T$. Growth is inversely proportional to $\mid \mathbf{\Sigma}_p \mid$, so that more stable estimates lead to faster growth.

6. If region growth is too slow, a switch is made to a higher order model (similarity to affine or affine to quadratic).

Termination with success occurs when the region expands to cover the entire area of overlap between the two images and the transformation error is sufficiently low (1.5 pixels, as determined empirically [7]). Termination with failure occurs when the initial error is too large or when the region is too unstable at the quadratic model to expand.

5. RESULTS AND DISCUSSION

An initial round of experimental results has been completed. In a large suite of image pairs taken from healthy eyes the algorithm succeeds xx% more often than our earlier algorithm [7]. An example result is shown in Fig. 5. In a smaller dataset of images from diseased eyes, the new algorithm has not yet failed to produce an alignment. The example shown in Fig. 3 is the most extreme case we've tried. The algorithm works in just a few seconds, on average.

The new registration algorithm succeeds because of the combination of two major ideas. The first is generating initial transformation hypotheses that are of insufficent complexity and only locally accurate. This gives the algorithm only a "foot-in-the-door" for a good transformation estimate, but, crucially, it avoids the need for a large number of initial correspondences. The second idea is the gradual extension and refinement of the transformation using the estimate's covariance matrix as a guide. The genesis of the technique can be seen in our recent work [9], but the current algorithm is novel in its ability to use a single correspondence and, most importantly, in the application of the covariance matrix. The ideas presented here can be applied to many other medical image registration problems.

6. REFERENCES

[1] J Roider and H. Laqua, "Laser coagulation of age-related macular degeneration," *Klinische Monatsblatter Fur Augenheilkunde*, vol. 206, pp. 428–437, June 1995.

[2] E.M. Kohner, I.M. Sleightholm, S.J. Aldington, R.C. Turner, and D.R. Matthews, "Microaneurysms in the development of diabetic retinopathy," *Diabetologia*, vol. 42, pp. 1107–1112, 1999.

[3] S. Fine, "Observations following laser treatment for choroidal neovascularization," *Archives of Ophthalmology*, vol. 106, pp. 1524–1525, 1988.

[4] T.M. Clark, W.R. Freeman, and M. Goldbaum, "Digital overlay of fluorescein angiograms and fundus images for treatment of subretinal neovascularization," *retina*, vol. 12, pp. 118–26, 1992.

[5] W.E. Hart and M.H Goldbaum, "Registering retinal images using automatically selected control point pairs," in *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, 1994, vol. 3, pp. 576–581.

[6] F. Zana and J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," *IEEE Trans. on Medical Imaging*, vol. 18, no. 5, pp. 419–428, 1999.

[7] A. Can, C.V. Stewart, B. Roysam, and H.L. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. on PAMI*, vol. 24, no. 3, 2002.

[8] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from live high-resolution retinal fundus images using direct exploratory algorithms," *IEEE Trans. on Info. Tech. for Biomedicine*, vol. 3, no. 2, pp. 125–138, 1999.

[9] H. Shen, C.V. Stewart, B. Roysam, G. Lin, and H.L. Tanenbaum, "Frame-rate spatial referencing based on invariant indexing and alignment with application to laser retinal surgery," in *Proc. CVPR*, 2001, vol. 1, pp. 79–86.

[10] Y. Chen and G.G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, vol. 10, no. 3, pp. 145–155, 1992.

[11] Charles V. Stewart, "Robust parameter estimation in computer vision," *SIAM Reviews*, vol. 41, no. 3, September 1999.

[12] Richard Hartley and Andrew Zisserman, *Multiple View Geometry*, Cambridge University Press, 2000.