This received the "best paper" award at the *1999 IEEE Conf. on Computer Vision and Pattern Recognition*

Robust Hierarchical Algorithm for Constructing a Mosaic from Images of the Curved Human Retina*

Ali Can[1] Charles V. Stewart[2] Badrinath Roysam[1]

[1]Dept. of Elec. Comp. and Sys. Eng.
Rensselaer Polytechnic Institute
Troy, NY 12180-3590
{cana,roysab}@rpi.edu

[2]Dept. of Computer Science
Rensselaer Polytechnic Institute
Troy, NY 12180-3590
stewart@cs.rpi.edu

Abstract

This paper describes computer vision algorithms to assist in retinal laser surgery, which is widely used to treat leading blindness causing conditions but only has a 50% success rate, mostly due to a lack of spatial mapping and reckoning capabilities in current instruments. The novel technique described here automatically constructs a composite (mosaic) image of the retina from a sequence of incomplete views. This mosaic will be useful to opthalmologists for both diagnosis and surgery. The new technique goes beyond published methods in both the medical and computer vision literatures because it is fully automated, models the patient-dependent curvature of the retina, handles large interframe motions, and does not require calibration. At the heart of the technique is a 12-parameter image transformation model derived by modeling the retina as a quadratic surface and assuming a weak perspective camera, and rigid motion. Estimating the parameters of this transformation model requires robustness to unmatchable image features and mismatches between features caused by large interframe motions. The described estimation technique is a hierarchy of models and methods: the initial match set is pruned based on a 0^{th} order transformation estimated using a similarity-weighted histogram; a 1^{st} order, affine transformation is estimated using the reduced match set and least-median of squares; and the final, 2^{nd} order, 12-parameter transformation is estimated using an M-estimator initialized from the 1st order results. Initial experimental results show the method to be robust and accurate in accounting for the unknown retinal curvature in a fully automatic manner, while preserving image details.

*The authors would like to thank Dr. Howard L. Tanenbaum of the Center for Sight, Albany, NY, Dr. James N. Turner of the Wadsworth Center for Laboratories and Research, NY State Dept. of Health, and Amitha Perera Dept. of Computer Science, Renssealer for their contributions to this work.

1 Introduction

Medical applications of computer vision have received increasing attention over the past several years. Many of these applications are in MRI and CT image analysis [20], while others are in robotic surgery [27], and in surgical planning and visualization [11]. One application area where the development of 3-D computer vision techniques has been less substantial, but where the need is great, is in opthalmology. This paper considers the particular problem of retinal laser surgery.

Retinal laser surgery, currently done manually, is a proven treatment for leading blindness-causing conditions [17, 22], including age-related macular degeneration [19], degenerative myopia, and diabetic retinopathy, affecting over 24 million people in the US alone [21, 31]. When diagnosing these conditions, the retina is imaged in two spectral channels. One channel images the retina in the visible spectrum, while the other images the layer beneath the retina in the near-infrared using a fluorescent dye (Indocyanine Green [14]). Using both types of imagery, physicians are able to delineate pathological regions for subsequent surgical laser photocoagulation. This surgical procedure must be done accurately, without damaging vital healthy tissue, without missing pathological tissue, and without overexposure [22]. Although laser retinal surgery is the most widely used treatment for the above conditions, the current success rate is below 50% for the first treatment, with a recurrence and/or persistence rate of about 50%. Each re-treatment also has a 50% failure rate, with visual recovery declining with successive treatments [31]. These failures can be traced to the manual nature of the surgery, and the lack of spatial reckoning aids in the clinical instrumentation. It is difficult for the physician to develop a complete view of the retina, to respond quickly to patient eye movements which can not be completely stopped during surgery, and to measure optical dosage quantitatively. The compelling need to minimize invasiveness implies that computer vision methods are preferable for addressing the above problems [29].

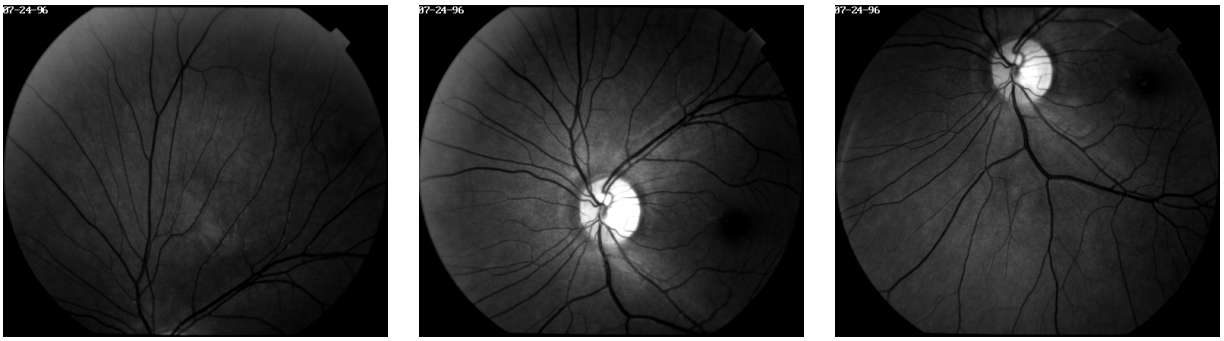

Figure 1: Three example 1024×1024 images of the retina surface showing relatively little overlap between them.

We are developing a two-phase computer vision system to assist physicians during diagnosis and retinal laser surgery. Phase I, coinciding with the physician's exploration for diagnosis, involves constructing a high-resolution wide-area mosaic of the retina. Phase II, coinciding with surgical treatment, involves real-time tracking of the projected position of the laser on the retina and comparing this against physician-delineated regions of surgical interest. Tracking results will be used to calculate dosimetry and to turn the laser on or off as appropriate. The focus of the current paper is on the mosaic constructed during Phase I and used as a spatial reference map during Phase II. The image matching technique developed for mosaic construction will also be used as part of a multistage, hierarchical tracking algorithm.

Although the problem of constructing a composite or "mosaic" image from image sequences has been studied heavily over the past decade in both the computer vision [15, 26, 25] and biomedical literatures [18, 5, 9], the problem of automatically constructing a high-resolution mosaic from a retinal image sequence raises two major issues requiring novel techniques:

- An appropriate inter-image transformation model must be found for mapping individual images onto the mosaic image frame. This model must account for the curved surface of the retina, the relatively general rigid motion possible, and the inherent lack of calibration because the optics of each individual retina changes the effective optics of the camera. We will demonstrate the need for a second-order transformation model which we derive here.

- A new technique is needed to estimate parameters of the transformation model. Because large eye movements are possible, there may be as little as 40% overlap between images. This implies that feature-based techniques [28, 30] should be used instead of intensity-based methods [3, 6], but more importantly it indicates that robustness to large numbers of mismatches and unmatchable features must be central to the estimation technique.

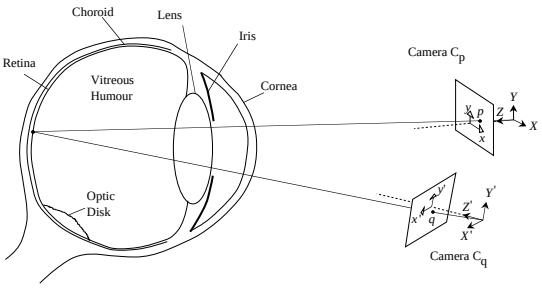

Figure 2: Two views of the retina.

Investigation of these issues forms the heart of the paper.

2 Image Transformation Model

The retinal surface is almost but not quite spherical, making a quadric or even a quadratic model, intuitively, a good approximation to its shape. During diagnosis and surgery, a patient's pupil is dilated and his/her forehead is held against a harness. Small shifts in head position are likely, inducing translations and rotations of the eye. Eye movements themselves, not significantly constrained during diagnosis or surgery, are almost exclusively rotational and can occur about two axes at rates of up to 180° per second. Significantly, neither axis of rotation is the optical axis. The fundus camera through which the retina is viewed is fixed and produces high-resolution, 1024×1024 images (Fig. 1). Together, the foregoing observations imply that the apparent motion of the retina should be modeled as a general rigid motion. (They also imply, however, that some components of the motion — rotation about the camera's optical axis in particular — will be small.) This generality suggests that the curved nature of the retina surface should be taken into account in the inter-image transformation model and that transformation models based on a planar surface [1, 6] will be insufficient. We confirm these intuitions by deriving a

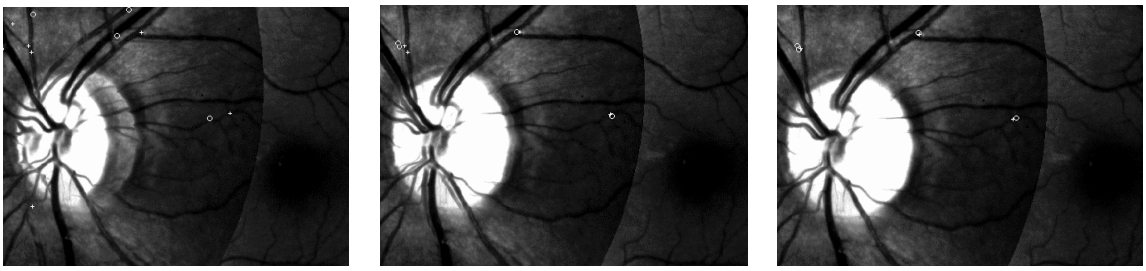

Figure 3: Examples showing visually the transformation error obtained using a 0th order — image translation only — model (left), the affine model (center), and the quadratic model (right).

transformation model for a quadratic surface imaged by a weak perspective camera and by comparing it to models for planar surfaces.

Consider two images $I_p(\mathbf{p})$ and $I_q(\mathbf{q})$ of a quadratic surface (the retina) taken by weak perspective cameras C_p and C_q (Fig. 2). Let 3×3 rotation matrix $\mathtt{R}$ and 3×1 vector $\mathbf{t}$ define the rigid transformation between camera viewpoints. The transformation may be due to movement of the surface. Let $(X,Y,Z)^T$ be a point on the surface in the coordinate system of camera C_p having the Z axis aligned with the optical axis, so that $\mathbf{p} = (x,y)^T = s(X,Y)^T$. Write the quadratic surface equation as

$$Z = (1,X,Y)\,\mathtt{B}\,(1,X,Y)^T, \tag{1}$$

with $\mathtt{B}$ a 3×3 symmetric parameter matrix. Also, let $(X',Y',Z')^T$ describe the same point as $(X,Y,Z)^T$ but in the coordinate system of C_q, so that $\mathbf{q} = (x',y')^T = s'(X',Y')^T$. Then, because of the rigid transformation

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} t_X \\ t_Y \end{pmatrix}$$

where the r_{ij} are components of $\mathtt{R}$. Substituting $(1,X,Y)\,\mathtt{B}\,(1,X,Y)^T$ for Z, then substituting x/s for X, y/s for Y, x'/s' for X' and y'/s' for Y', and finally manipulating the equation yields

$$\mathbf{q} = \Theta\,\mathbf{x}(\mathbf{p}), \tag{2}$$

where $\mathbf{x}(\mathbf{p}) = (1, x, y, x^2, xy, y^2)^T$ and Θ is a 2×6 parameter matrix. We only estimate the 12 parameters of Θ rather than extracting the unknown quadratic surface, camera, and motion parameters represented by Θ. Equation 2 generalizes the 2D affine transformation model induced by the rigid motion of a planar surface [6], which is given by:

$$\mathbf{q} = \mathtt{A}\mathbf{p} + \mathbf{t}. \tag{3}$$

Notice that (2) is the second-order Taylor series expansion of the general image transformation equation, while the affine model (3) is the first-order expansion. The instantaneous motion of a planar surface under a full-perspective model induces an 8-parameter model, whereas that of a quadratic surface induces a 17-parameter, non-linear model.

The two planar models and the new quadratic models may be compared both numerically and visually. To do so, we manually locate and match features (bifurcations of the retinal vasculature) in pairs of images. We then find the image transformation parameters for each model minimizing the sum-squared error between a transformed feature and its match, using the standard deviation of the residual error to measure transformation model accuracy. Accuracies in the range 4-5 pixels are typical for the 6- and 8-parameter models induced by planar surface motion whereas accuracies of around 1.1 pixels are typical of transformations induced by quadratic surface motion. This is particularly encouraging because error in manual feature position localization is on the order of a pixel. We therefore choose the 12-parameter model, which we refer to as the "quadratic model", for constructing retinal mosaics. This choice is confirmed by visual comparison between the affine and quadratic models shown in Fig. 3(center) and (right). The advantage of the quadratic model can clearly be seen near the white center (the optic disk). The transformation estimated under simple image translation is also shown. The significance of this will be explained in the next section.

3 Estimating the Transformation

The next step is developing a technique to estimate the quadratic model parameters for any pair of image frames. The major challenge is handling potentially small overlap between images. This necessitates use of feature-based matching as the basis for estimating the transformation and suggests the need for substantial robustness to unmatched and mismatched features.

3.1 Detecting Vasculature Bifurcations

Looking at the images in Fig. 1, the most prominent features are the blood vessels and their bifurcations/crossovers. Other image regions are usually relatively featureless. Additional features appear in pathological retinas, but vascular

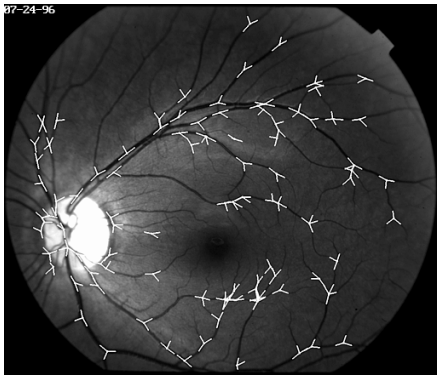

Figure 4: Landmarks detected using the recursive vasculature tracing algorithm.

structure usually continues to be prominent. We therefore detect bifurcations and match them as the basis for estimating the inter-image transformation.[1]

We have developed a recursive vasculature tracing algorithm to map out the blood vessel structure and detect bifurcations ("landmarks") in an image. The algorithm, which has been described in detail elsewhere [8], starts by analyzing the image along a coarse grid to do 1D edge detection and gather image intensity statistics. Seed points for tracing are identified and prioritized from a detailed image analysis around detected edges. Subsequent recursive tracing is straightforward and rapid, with special case checks to detect bifurcations. An example result is shown in Figure 4. The algorithm runs at frame rate on a dual processor SGI.

3.2 Hierarchical Estimation Algorithm

Matching bifurcations and estimating the quadratic transformation requires special care. The bifurcations are not particularly distinguishable from each other locally because many have similar vessel orientations and backgrounds. Therefore, the large motion between frames with resulting small image overlap implies that there will be many different possible matches for each bifurcation, but many bifurcations may have no correct matches at all. This makes the matching and transformation estimation problem appear more difficult than related robust matching problems such as in fundamental matrix estimation [28, 30].

To motivate the estimation technique, we revisit the discussion and results of Sec. 2. First, the retinal surface, while approximately quadratic, is roughly planar over large regions. This explains why the affine transformation model yields small error, and suggests that an affine model, involving only 6 parameters and requiring only 3 correspondences, could be estimated as a preliminary step to estimating the full quadratic model. Second, the lack of rotation about the (eye's) optical axis implies that the apparent rotation in the image will not be substantial: rotations about the other two axes, each parallel to the image plane, produce apparent motions that are difficult to distinguish from image translation except over broad image regions. Third, translational motion of the head, especially along the optical axis is likely to be small enough that convergent and divergent motions in the image are insignificant. These second and third points suggest that a 0^{th} order model involving only image translation may be used as an even cruder initial approximation to the transformation (Fig. 3(a)). Fourth and finally, the features are relatively sparse and the features sets are relatively small, involving 30-50 bifurcations. This makes practical the techniques described below.

We have developed a hierarchical matching and transformation estimation algorithm based on these observations. Each level of the hierarchy involves a different transformation model and estimation technique, and the results of one level are used as initial conditions for subsequent levels.

0th Order, Translation: Initially, just a translation vector, $\mathbf{t}_0$, is estimated. This represents the 0th order transformation model

$$T_0(\mathbf{p};\mathbf{t}_0) = \mathbf{p} + \mathbf{t}_0 \tag{4}$$

Let the initial vasculature bifurcation sets from the two images be denoted $\mathbf{P}$ and $\mathbf{Q}$, with N_p and N_q points respectively. Form a 2D weighted histogram of translation vectors with large histogram bin sizes (e.g. 15-20 pixels on a side), and establish an initial correspondence set $\mathbf{C}_0 = P \times Q$. Restrictions could be placed on the matches $\mathbf{C}_0$, but thus far this has not been necessary. For each $(\mathbf{p}_i, \mathbf{q}_j) \in \mathbf{C}_0$, calculate $\mathbf{t}_{i,j} = \mathbf{q}_j - \mathbf{p}_i$ and enter a weighted vote in the histogram bin containing $\mathbf{t}_{i,j}$. The weight is a similarity measure on the matches based on similarity in intersection angles and vessel thickness. Figure 5 shows an example histogram. The peak of this histogram, after low-pass filtering, provides the estimate $\hat{\mathbf{t}}_0$. A new correspondence set $\mathbf{C}_1$ is formed by all matches such that $\mathbf{t}_{i,j}$ is within twice the histogram bin width of $\hat{\mathbf{t}}_0$. Some features will have several correspondences in $\mathbf{C}_1$, while others will have none.[2]

1st Order, Affine Model: The next level of the hierarchy estimates an affine transformation

$$T_1(\mathbf{p};\mathtt{A},\mathbf{t}_1) = \mathtt{A}\mathbf{p} + \mathbf{t}_1$$

using a minor variation on least-median of squares algorithm (LMS) [23]. Let $\mathbf{P}_1 \subseteq \mathbf{P}$ contain the features from I_p having at least one match in $\mathbf{C}_1$, and for each $\mathbf{p} \in \mathbf{P}_1$ let $\mathbf{C}_1(\mathbf{p}) = \{\mathbf{q} \mid (\mathbf{p},\mathbf{q}) \in \mathbf{C}_1\}$. Note that if the overlap between

[1] In ongoing work we are studying the possibility of detecting and matching pathologies such as drusen, lipid exudates, angioid streaks, hemorrhages, laser scars, and cotton-wool spots to augment matching of bifurcations.

[2] Points $\mathbf{p}$ are considered to be coming from the image to be incorporated into the mosaic, hence the asymmetric treatment of point sets $\mathbf{P}$ and $\mathbf{Q}$ in matching.

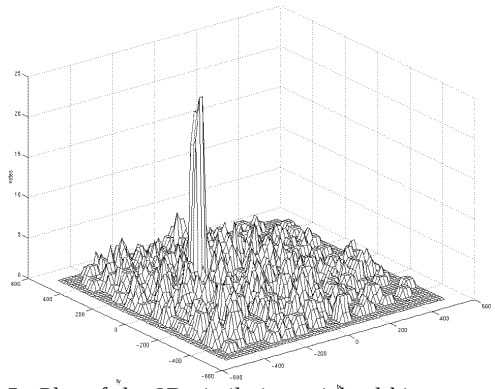

Figure 5: Plot of the 2D similarity-weighted histogram for one image pair.

images is small, $\mathbf{P}_1$ will be much smaller than $\mathbf{P}$. The LMS estimate of the affine parameters is

$$(\hat{\mathtt{A}}, \hat{\mathbf{t}}_1) = \underset{\mathtt{A},\mathbf{t}_1}{\operatorname{argmin}} \underset{\mathbf{p}\in\mathbf{P}_1}{\operatorname{median}} \min_{\mathbf{q}\in\mathbf{C}_1(\mathbf{p})} \|\mathbf{q} - \mathtt{A}\mathbf{p} - \mathbf{t}_1\|^2. \quad (5)$$

In other words, the objective function for a given $(\mathtt{A}, \mathbf{t}_1)$ is calculated by finding the minimum distance match for each transformed $\mathbf{p} \in \mathbf{P}_1$ and then taking the median of the resulting squared (error) distances. This differs from other well-known uses of LMS, e.g. in estimating the fundamental matrix [28, 30], because uniqueness in the correspondence set is not (and should not be) enforced yet. Minimization is accomplished through a random sampling technique [10, 23]. To form each sample, a triple of points is selected randomly from $\mathbf{P}_1$ and then a match for each point $\mathbf{p}$ in the triple is randomly chosen from $\mathbf{C}_1(\mathbf{p})$. After minimization, a robust scale value, $\hat{\sigma}_1$, is calculated from $(\hat{\mathtt{A}}, \hat{\mathbf{t}}_1)$ using the median absolute deviation (MAD) scale estimator [24]. Note that both uncertainty in feature position and modeling error in using the affine transformation contribute to the magnitude of $\hat{\sigma}_1$.

2nd Order, Quadratic Model: The final level of the hierarchy estimates the quadratic transformation

$$T_2(\mathbf{p}; \Theta) = \Theta\mathbf{x}(\mathbf{p})$$

(see Eqn. 2 and discussion) using an M-estimator [12]. We describe first a straightforward instantiation, and then show several important modifications. For each $\mathbf{p_i} \in \mathbf{P}_1$, let

$$\mathbf{q}_i = \underset{\mathbf{q}\in C_1(\mathbf{p_i})}{\operatorname{argmin}} \|\mathbf{q} - \hat{\mathtt{A}}\mathbf{p} - \hat{\mathbf{t}}_1\|^2,$$

i.e. $\mathbf{q}_i$ is the best match for $\mathbf{p}_i$ based on the estimated affine transformation. Then the M-estimate of Θ is

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{\mathbf{p}_i\in\mathbf{P}_1} \rho(\|\mathbf{q}_i - \Theta\mathbf{x}(\mathbf{p}_i)\|/\hat{\sigma}), \quad (6)$$

where ρ is a "robust loss function" which grows subquadratically, and $\hat{\sigma}$ is a scale estimate. Here we use the Tukey biweight function [4]:

$$\rho(u) = \begin{cases} \frac{a^2}{6}\left[1 - \left(1 - \left(\frac{u}{a}\right)^2\right)^3\right] & |u| \le a \\ \frac{a^2}{6} & |u| > a \end{cases}$$

where $u = \|\mathbf{q} - \Theta\mathbf{x}(\mathbf{p})\|/\hat{\sigma}$ is a "scale normalized residual." (Typically $a \approx 4.0$ [13].) Equation 6 may be solved using iteratively-reweighted least-squares (IRLS) [13], with weight function $w(u) = \rho'(u)/u$, in our case

$$w(u) = \begin{cases} \left[1 - \left(\frac{u}{a}\right)^2\right]^2 & |u| \le a \\ 0 & |u| > a. \end{cases}$$

The value u_i for each match in each iteration is calculated using the estimate of Θ obtained in the previous iteration. A robust starting point for IRLS is crucial, as has been confirmed experimentally in vision applications [28].

Several modifications of the estimation equation (6) and IRLS search technique are important here:

1. The M-estimator is initialized from the affine estimate $(\hat{\mathtt{A}}, \hat{\mathbf{t}}_1)$ and scale $\hat{\sigma}_1$. These are used to compute initial scale normalized residuals

   $$u_i = \|\mathbf{q}_i - \hat{\mathtt{A}}\mathbf{p}_i - \hat{\mathbf{t}}_1\|/\hat{\sigma}_1,$$

   and these are used to compute initial weights to be used in IRLS.

2. A new (MAD) scale, $\hat{\sigma}_2$, is estimated from the residuals of the first IRLS estimate of Θ and then fixed for the remaining iterations.

3. Because $w(u) = 0$ for residuals (error distances) greater than about $4\hat{\sigma}_2$, there is no need to restrict the match set in (6). The entire original set $\mathbf{C}_0$ may be used, since matches with large error distances simply contribute zero weight. (In practice, loose restrictions are placed purely for computational reasons.) This allows recovery from earlier mistakes in reducing the match set.

4. The robust weights $w(u)$ are augmented in IRLS by a correspondence similarity measure and, when more than one match for a given $\mathbf{p} \in \mathbf{P}$ has non-zero weight, these weights are normalized. For example, if $(\mathbf{p}, \mathbf{q})$ and $(\mathbf{p}, \mathbf{q}')$ are two matches for $\mathbf{p}$, with robust weights w and w' and similarity measures s and s', then the actual IRLS weights will be:

   $$w^* = sw\frac{sw}{sw + s'w'}; \qquad w'^* = s'w'\frac{s'w'}{sw + s'w'}$$

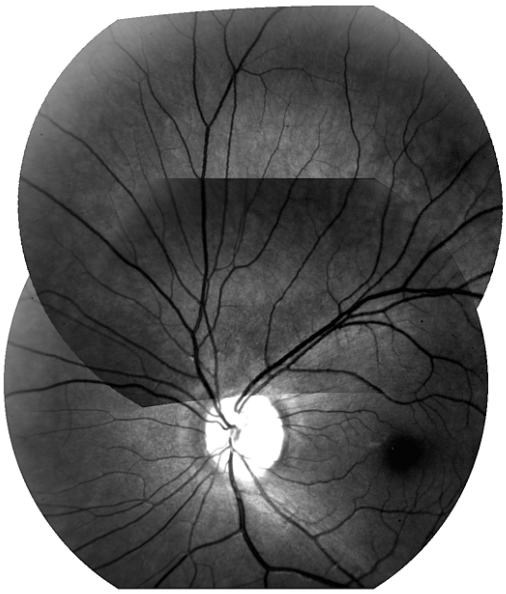

Figure 6: The combination of two images having relatively small overlap based on the estimated quadratic transformation. Nonlinear warping from the quadratic model can be seen at the edge of the top frame.

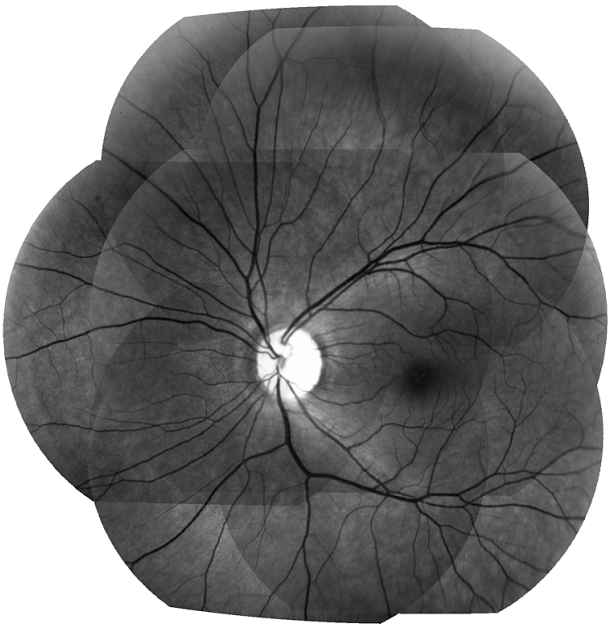

Figure 7: A mosaic combining 10 images each 1024×1024 pixels using the quadratic transformation model and hierarchical estimation technique.

In effect, the last two modifications allow decisions about correspondence to be deferred until the final transformation is estimated. This differs significantly from other robust matching and estimation algorithms [28, 30], and is a major reason why the algorithm is successful. Correct correspondences are determined to be those having non-zero weights. More than one such correspondence for any $\mathbf{p}$ indicates a lingering ambiguity, which is usually caused by merging of blood vessel intersections into a single feature in one image but not in the other. These are resolved and final correspondences for all features are established by minimizing normalized SSD measures between a template centered on $\mathbf{p}$ in I_p and regions surrounding $T_2(\mathbf{p}_i, \hat{\Theta})$ in I_q. Final estimation of Θ is based on these correspondences.

An example result showing one image transformed onto a second based on the estimated quadratic transformation is shown in Fig. 6.

4 Constructing the Complete Mosaic

The interimage transform estimation technique forms the core of the mosaic construction algorithm. In the retinal surgery application, mosaic construction will be done online and interactively, with a physician specifying images to be added to the mosaic one frame at a time. This ensures complete coverage and use of good quality images. The dataset shown here was aquired under similar circumstances.

Given an image, I_p, to be added to the mosaic image (I_q in the notation that has been used thus far), the first step is estimation of the quadratic transformation and especially the final correspondence set. Next, the intensities in I_p are normalized. Then, the region of I_q covered by the transformation of I_p is determined. For each pixel $\mathbf{q}$ in this region, $\mathbf{q}$ is inverse mapped onto $\mathbf{p}$, the intensity is interpolated in I_p and, back in the mosaic frame, the resulting value is averaged with pixels from other images that map onto $\mathbf{q}$. Since inverting the quadratic transform directly is difficult, a quadratic transform is estimated from the final correspondences between I_p and the mosaic frame with roles of $\mathbf{p}$ and $\mathbf{q}$ reversed. This is then used for the inverse mapping. Incorrect transform estimates may be identified by comparing the (transformed) vasculature structure of I_p output by the recursive tracing algorithm (Sec. 3.1) to that of the mosaic using a digital distance map [7]. An example mosaic constructed from 10 images is shown in Figure 7.

5 Discussion and Conclusions

We have presented a quadratic image transformation model and a robust, hierarchical technique for estimating the transformation between two images of the retinal surface. These form the core of a method for constructing a wide-area, high-resolution mosaic of the human retina, improving on the mostly 2-D methods developed to date for this application [29, 5], and going beyond current mosaicing methods in the computer vision literature [15, 26, 25]. Improvements to our technique are still possible, however, and we are currently considering indexing methods [16] as a complement to the first two levels of the hierarchy.

High quality retinal mosaics such as those shown in Fig. 7 are directly valuable for diagnosing a variety of dis-

eases, especially those involving the retinal periphery [18]. Other intended uses of these mosaics as part of our long-term goal of developing a two-phase computer vision system to assist physicians during diagnosis and retinal laser surgery include: treatment planning, on-line treatment monitoring, spatial dosimetry, error detection, alarms and safety shutoffs when the laser strays from the intended target, change analysis between patient visits, and even virtual reality tools for surgical simulation. The mosaic construction algorithm, in modified form, can also enhance emerging alternatives to classic laser surgery [2] and be used to detect mismatches between small image regions to flag retinal bumps and detachments.

References

[1] G. Adiv. Determining 3-d motion and structure from optical flow generated by several moving objects. *IEEE Trans. on PAMI*, 7(4):384–401, 1985.

[2] S. Asrani, S. Zou, G. D'Anna, S. Lutty, S. Vinores, M. Goldberg, and R. Zeimer. Feasibility of laser-targeted photoocclusion of the choriocapillary layer in rats. *Invest. Ophthalmol. Vis. Sci.*, 38(13):2702–2710, Dec 1997.

[3] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *Int. J. of Computer Vision*, 12(1):43–77, February 1994.

[4] A. E. Beaton and J. W. Tukey. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16:147–185, 1974.

[5] D. E. Becker, A. Can, H. L. Tanenbaum, J. N. Turner, and B. Roysam. Image processing algorithms for retinal montage synthesis, mapping, and real-time location determination. *IEEE Trans. on Biomed. Eng.*, 45(1), Jan. 1998.

[6] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Proc. 2nd ECCV*, pages 237–252, 1992.

[7] G. Borgefors. Distance transformations in digital images. *CVGIP*, 34(3):344–371, June 1986.

[8] A. Can, 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 Biomed. Eng.*, 1999. Accepted, to appear.

[9] P. Dani and S. Chaudhuri. Automated assembling of images - image montage preparation. *Pattern Recognition*, 28(1):431–445, March 1995.

[10] M. A. Fischler and R. C. Bolles. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *CACM*, 24:381–395, 1981.

[11] W. Grimson, T. Lozano-Perez, W. Wells, III, G. Ettinger, S. White, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery and enhanced reality visualization. *IEEE Trans. on Med. Img.*, 15, 1996.

[12] F. R. Hampel, P. J. Rousseeuw, E. Ronchetti, and W. A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 1986.

[13] P. W. Holland and R. E. Welsch. Robust regression using iteratively reweighted least-squares. *Commun. Statist.-Theor. Meth.*, A6:813–827, 1977.

[14] L. Hyvarinen and R. W. Flower. Indocyanine green fluorescence angiography. *Acta Ophthalmologica*, 58(4):528–538, 1980.

[15] M. Irani, P. Anandan, and S. Hsu. Mosaic based representations of video sequences and their applications. In *Proc. IEEE Int. Conf. on Computer Vision*, pages 605–611, 1995.

[16] D. Jacobs. Matching 3-d models to 2-d images. *Int. J. of Computer Vision*, 21(1-2):123–153, January 1997.

[17] J. M. Krauss and C. A. Puliafito. Lasers in ophthalmology. *Lasers In Surgery And Medicine*, 17:102–159, 1995.

[18] A. A. Mahurkar, M. A. Vivino, B. L. Trus, E. M. Kuehl, M. B. Datiles, and M. I. Kaiser-Kupfer. Constructing retinal fundus photomontages. *Investigative Ophthalmology and Visual Science*, 37(8):1675–1683, July 1996.

[19] R. Murphy. Age-related macular degeneration. *Ophthalmology*, 9:696–971, 1986.

[20] W. Niessen, K. Vincken, J. Weickert, and M. Viergever. Three dimensional MR brain segmentation. In *Proc. IEEE Int. Conf. on Computer Vision*, pages 53–58, 1998.

[21] Research to Prevent Blindness Foundation. RPB: Research to prevent blindness annual report. 598 Madison Avenue, New York, NY 10022, 1992.

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

[23] P. J. Rousseeuw. Least median of squares regression. *J. Amer. Stat. Assoc.*, 79:871–880, 1984.

[24] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection.* John Wiley & Sons, 1987.

[25] H. Sawhney, S. Hsu, and R. Kumar. Robust video mosaicing through topology inference and local to global alignment. In *Proc. 5th ECCV*, volume II, pages 103–119, 1998.

[26] R. Szeliski. Video mosaics for virtual environments. *IEEE CGA*, 16(2):22–30, 1996.

[27] R. Taylor et al. An image-directed robotic system for precise orthopedic-surgery. *IEEE Trans. on Robotics and Automation*, 10(3):261–275, 1994.

[28] P. Torr and D. Murray. The development and comparison of robust methods for estimating the fundamental matrix. *Int. J. of Computer Vision*, 24(3):271–300, 1997.

[29] C. H. Wright, R. D. Ferguson, H. G. Rylander III, A. J. Welch, and S. F. Barrett. Hybrid approach to retinal tracking and laser aiming for photocoagulation. *Journal of Biomedical Optics*, 2(2), 195–203 1997.

[30] Z. Zhang, R. Deriche, O. Faugeras, and Q. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, 1995.

[31] I. E. Zimmergaller, N. M. Bressler, and S. B. Bressler. Treatment of choroidal neovascularization — updated information from recent macular photocoagulation study group reports. *International Ophthalmology Clinics*, 35:37–57, 1995.