Frame-Rate Spatial Referencing Based on Invariant Indexing and Alignment with Application to On-Line Retinal Image Registration*

Hong Shen[1]   Charles V. Stewart[1]   Badrinath Roysam[1]
Gang Lin[1]   Howard L. Tanenbaum[2]

[1]Rensselaer Polytechnic Institute
Troy, New York 12180–3590
stewart@cs.rpi.edu,

[2]The Center for Sight
349 Northern Blvd.
Albany, NY 12204

April 23, 2002

Abstract

This paper describes an algorithm to continually and accurately estimate the absolute location of a diagnostic or surgical tool (such as a laser) pointed at the human retina, from a series of image frames. We treat the problem as a registration problem, using diagnostic images to build a spatial map of the retina and then registering each on-line against this map. Since the image location where the laser strikes the retina is easily found, this registration determines the position of the laser in the global coordinate system defined by the spatial map. For each on-line image, the algorithm computes similarity invariants, locally valid despite the curved nature of the retina, from constellations of vascular landmarks. These are detected using a high-speed algorithm that iteratively traces the blood vessel structure. Invariant indexing establishes initial correspondences between landmarks from the on-line image and landmarks stored in the spatial map. Robust alignment and verification steps extend the similarity transformation computed from these initial correspondences to a global, high-order transformation. In initial experimentation, the method has achieved 100% success on 1024×1024 retina images. With a version of the tracing algorithm optimized for speed on 512×512 images, the computation time is only 51 milliseconds per image on a 900MHz Pentium III processor and a 97% success rate is achieved. The median registration error in either case is about 1 pixel.

Index terms: retinal laser surgery, registration, invariant indexing, robust estimation, retinal imaging, ophthalmic image processing, biomedical image processing.

---

*Hong Shen is now will Siemens Corporation, Princeton, NJ. An earlier version of this work was presented at the *IEEE Conference on Computer Vision and Pattern Recognition in December 2001.*

1 Introduction

Many techniques used in the diagnosis and treatment of diseases of the retina could be greatly improved by using automated, image-based tools that determine the locations of selected spots on the retina surface using clinical image sequences. We call this capability *spatial referencing*. The reference coordinate frame is a *spatial map* of the retina, pre-computed from off-line diagnostic images [8, 9, 10, 11] (Figure 1). On-line spatial referencing must be fast enough to enable real-time control decisions that are the key to developing diagnostic and treatment aids.

An important application for spatial referencing is laser retinal surgery.[1] This is the only proven treatment for leading blindness-causing conditions [17, 21, 22, 23], including age-related macular degeneration, degenerative myopia, and diabetic retinopathy, affecting over 24 million people in the US alone [13, 24, 25, 32], but still has only a 50% success rate. A major cause of failure is the lack of spatial information during surgery: there is no quantitative accounting of the energy delivered anywhere on the retina, and there are no automatic alarms or safety shut-offs when the laser is mis-aimed due to eye movement. These could be provided by detecting the location of the laser in surgical images of the retina in real-time, using the spatial referencing technique to transform this location onto the spatial map, and comparing this location with a physician-delineated treatment area.

One approach to spatial referencing is to track a retinal image region specified by a physician, starting from a known position [1, 19]. Unfortunately, frequent, large-scale changes between images caused by saccadic eye movements, blinking, and glare mean that incremental methods can only work for relatively short durations. What's needed is the capability of locating the laser spot absolutely, rapidly, and without reference to prior estimates. Such a capability could certainly be combined with tracking for even higher speeds as described in earlier work by Becker [2], but the capability itself is paramount.

---

[1]Laser surgery on the retina — the back of the eye — is completely different from the now-common laser surgery for lens correction.
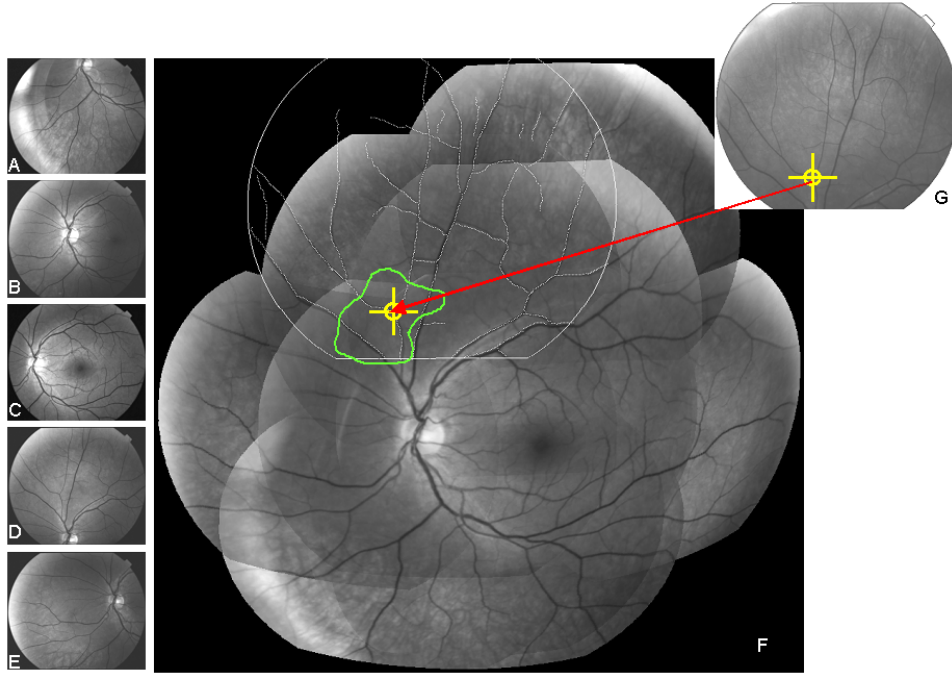
Figure 1: An illustration of spatial referencing. Diagnostic images shown in A-E can be combined in a mosaic that forms a map of the retina (F). The surgeon would delineate a region of surgical interest on this map. An example on-line image is shown at the upper right (G) and the location where a surgical laser beam would strike the retinal surface is illustrated using cross-hairs. The transformation of the on-line image onto the spatial map would determine the position of the laser on the spatial map and is therefore what must be computed in solving the spatial referencing problem. This would determine whether or not the laser is aimed within the treatment region drawn by the ophthalmologist.

Thus, the goal of the research described in this paper is to develop a method capable of high-speed spatial referencing on modest computing hardware without use of incremental tracking. The required speed is at least standard frame rates — 30 images per second. The key idea is the use of invariant indexing on localized sets (constellations) of retinal image landmarks — branching and cross-over points in the retinal vasculature (Figure 2). Invariants may be used as a signature vector on a set of landmarks in one image that uniquely identifies the same set of points in different images. Vectors of invariants may be precomputed — in the diagnostic images, for example — and stored in a data structure for rapid on-line look-up (indexing). Correspondences between landmarks, established by indexing, may be used to determine an initial transformation between the on-line image and the spatial map (Figure 3).
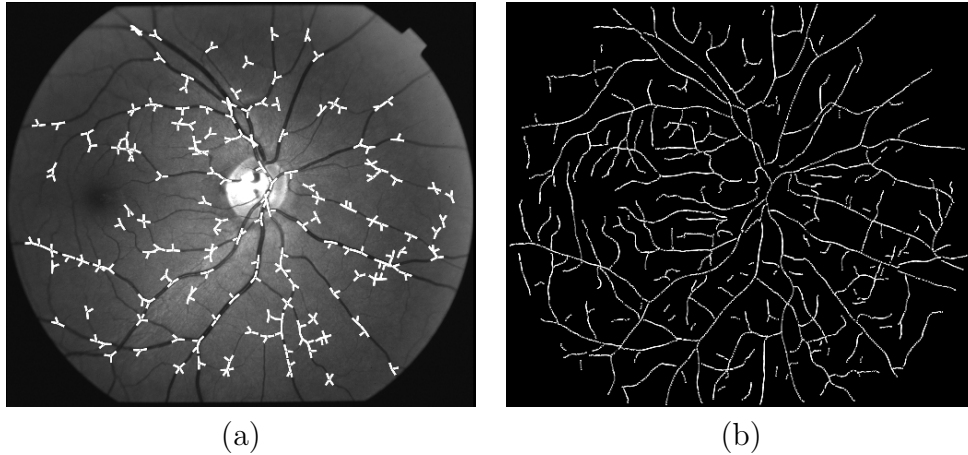
<center>(a)   (b)</center>

Figure 2: Illustrating the results of iterative tracing and vascular landmark detection: (a) the vascular landmarks drawn on top of the image, and (b) the network of traced vascular centerline points. Constellations of landmarks are used in indexing, while the centerline points are used in verification and refinement.

This transformation may then be verified, extended and refined to cover the entire on-line image. In effect, this approach turns spatial referencing into an identification, alignment and verification problem reminiscent of techniques from object recognition [16, 26, 33].

Several issues must be addressed to make this approach effective and efficient: (1) the choice of transformation models and invariants, (2) high-speed feature extraction, (3) pre-computation of invariants and the spatial map from diagnostic images, and (4) the actual on-line indexing, alignment and verification procedure. The spatial map includes the invariant database, the diagnostic images and associated digital distance maps, and inter-image transformations between diagnostic images.

2   Transformation Models

Three transformation models are used in the indexing, refinement and verification procedures. The final and most accurate model is a quadratic model derived in [8, 11]. Let $\mathbf{x} = (x, y)^T$ be the coordinates of a point location in one image, I_p, and let $\mathbf{u} = (u, v)^T$ be the coordinates of the same location on the retina in a second image, I_q. Then, the

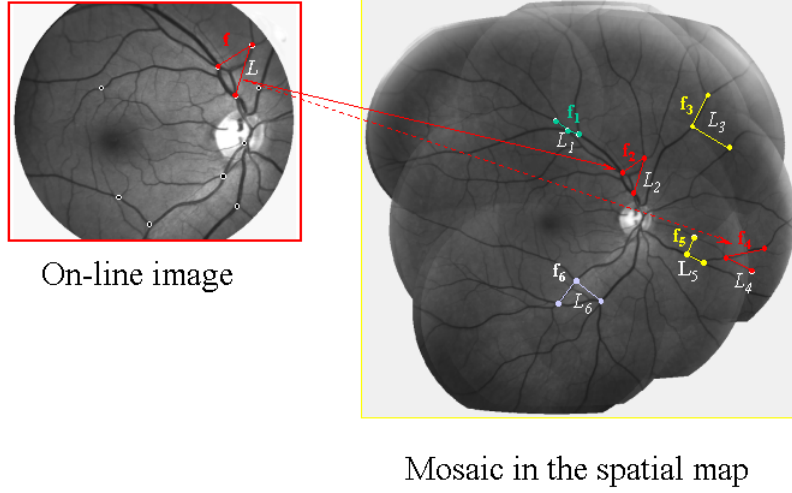<center>3</center>

On-line image

Mosaic in the spatial map

Figure 3: Illustrating indexing. A vector, $\mathbf{f}$, of quasi-invariants is computed from a constellation of landmarks, L, in an on-line image. The database of vectors pre-computed from off-line images and stored in the spatial map is indexed to quickly find approximately matching vectors ($\mathbf{f}_2$ and $\mathbf{f}_4$) and their associated landmark constellations (L_2 and L_4). Correspondences between landmarks in the constellation are used to generate transformation estimates. These are locally tested and rejected (the match between L and L_4), or locally tested, refined and extended to cover the entire on-line image (the match between L and L_2).

12-parameter quadratic transformation mapping $\mathbf{x}$ onto $\mathbf{u}$ is

$$\mathbf{u} = \mathbf{M}_Q(\mathbf{x}; \boldsymbol{\Theta}) = \boldsymbol{\Theta} \cdot \mathbf{X}(\mathbf{x}) \tag{1}$$

where $\mathbf{X}(\mathbf{x}) = (x^2, xy, y^2, x, y, 1)^T$ and $\boldsymbol{\Theta}$ is a 2×6 matrix of transformation parameters. This model is accurate to less than a pixel on 1024×1024 images and is the transformation that must ultimately be estimated in solving the spatial referencing problem.

The other two models are approximations to this quadratic model — two-dimensional affine and similarity transformations. These have 6 and 4 degrees of freedom and average 2.5 and 3.1 pixels of error respectively [11]. They are reasonably accurate because of the restricted viewpoint on the eye imposed by the iris and the fundus camera [28].

3   Similarity Invariants

In designing the invariants for spatial referencing, three issues must be addressed: the choice of transformation model, the number of landmarks used in forming an invariant and, based on these, the actual invariant signature vector.

Our choice of models is the similarity transformation. The quadratic transformation does not yield any invariants. (It is not closed under composition, implying that it does not form a transformation group and therefore has no invariants.) The similarity transformation is chosen over the affine because the slight increase accuracy in using the affine model is not worth the extra two degrees of freedom (see [28]). The resulting invariants are really quasi-invariants because they change slowly with changes in viewpoint [5].

Similarity invariants are computed from pairs of landmarks and from triples of landmarks. These pairs and triples are called "constellations". The distance between any two landmarks in a constellation must be less than 20% of the image width. This empirical threshold allows formation of enough constellations while retaining sufficient accuracy in the use of the similarity model. Increasing the number of landmarks beyond three results in the formation of too few constellations, especially on the image periphery [28]. The advantage of using triples in a constellation is that more distinctive invariants may be computed. The advantage of using pairs is that fewer constellations are formed (by a factor of 4) and any particular constellation is less likely to be missed by missing a landmark.

The invariant vectors are computed from the positions and orientations of the landmarks forming the constellations. These result in what we refer to as "quasi-invariant feature vectors" (QIFVs). A five-dimensional QIFV is computed from triples of landmarks (see Figure 4). Let the landmark locations be $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, ordered so that the segment $(\mathbf{x}_1, \mathbf{x}_2)$ is the longest and $(\mathbf{x}_1, \mathbf{x}_3)$ is the shortest. All components of the QIFV are computed by mapping the landmarks into a canonical similarity basis, placing $\mathbf{x}_1$ at the origin, $(0,0)^T$, and $\mathbf{x}_2$ at location $(0,1)^T$. The position of $\mathbf{x}_3$ when transformed into this canonical coordinate system forms the first two components, α and β, of the QIFV. (The invariance of α and β
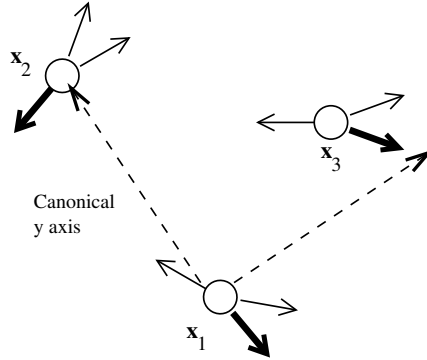
Figure 4: Forming the 5-component similarity quasi-invariant feature vector (QIFV) from 3 points. Point $\mathbf{x}_3$ maps to $(\alpha, \beta)^T$ in the canonical basis formed by placing $\mathbf{x}_1$ at the origin, $(0,0)^T$, and $\mathbf{x}_2$ at $(0,1)^T$ (forming the y-axis and the scale). α and β are the first two components of the QIFV. The directions shown in bold from each landmark, measured in the canonical basis, form the other three components.

is easily seen by applying a similarity transformation to each $\mathbf{x}_i$, and then reapplying the canonical basis mapping.) The other three invariants are formed by one angle from each of the three landmarks, mapped into the canonical coordinate system. Letting the range of the mapped angles be $(-\pi, \pi]$ with respect to the canonical x-axis, the most negative angle is chosen for each landmark.

In the computation of a QIFV from a pair of landmarks, $\mathbf{x}_1$ and $\mathbf{x}_2$, the relative positions are used to form the similarity basis. The orientations of each of the six blood vessels emerging from the landmarks, when mapped into this basis, form the invariants.

4 The Spatial Map

The spatial map contains all the information that can be computed off-line from the diagnostic images to support fast, on-line spatial referencing. The use of similarity invariants drives the design of the spatial map. Fundamentally, the similarity transformation model is only accurate in individual images and most accurate near the center of these images. This means that the invariants should be computed from individual diagnostic images, not from combined mosaic images (despite the illustrative examples in Figures 1 and 3), since doing

so would create distortions in quasi-invariant feature vectors (QIFVs) computed based on the similarity model. The diagnostic images are represented individually in the spatial map, with pre-computed and known transformations between them to associate the QIFVs from the same constellation in different images.

Here is a summary of the spatial map computation and storage given N diagnostic images.

1. Locate vascular centerline points and landmarks as shown in Figure 2 using the tracing algorithm described below.

2. Estimate the quadratic transformation between each pair of images [11, 10].

3. Find landmark constellations and compute invariants (QIFVs):

   (a) In each image form the landmarks into constellations of two or three, depending on which QIFVs are being used.

   (b) Use the estimated quadratic transformations to find the same landmark constellation in different (overlapping) images. These constellations are formed into "constellation groups".

   (c) Compute invariants for each landmark constellation and store the QIFVs in a k-d tree (as in [3]). Each constellation also stores a pointer to its constellation group.

4. For each image compute a digital distance map [6]. This map has the same dimensions as the image and stores two pieces of information at each pixel: the distance to the nearest vascular centerline point (Figure 2), and a pointer to this point.

Details on the tracing algorithm and the storage of constellations follow.

Each off-line image is processed to extract blood vessel centerlines and to detect landmarks — branches and cross-over points of blood vessels (Figure 2). This uses an exploratory algorithm that iteratively traces the vasculature structure starting from seed points detected by analyzing the image along a grid of 1-d cross-sections [7, 29]. The algorithm repeatedly detects anti-parallel edges forming a blood vessel boundary, places a blood vessel centerline

point midway between the edge locations, and steps parallel to the direction of the edges. Landmark locations are determined from the meeting and crossing of these centerline traces, and then refined to subpixel accuracy [31].

Once all landmark constellations are detected in all diagnostic images and then grouped between images, their QIFVs are stored in a database for fast access during on-line indexing. The QIFV's for each constellation are stored separately rather than as a summary description for the constellation group because (1) there may be too few samples for each constellation group to compute accurate summary statistics on the QIFVs, (2) a mean vector and covariance matrix for each constellation group may not accurately represent the distribution of QIFVs, and (3) on-line, the search for the closest QIFV is simpler when each constellation is represented separately.

5   On-line Spatial Referencing

On-line, during surgery, the computational goal is to estimate the transformation between each image and the spatial map as quickly and accurately as possible. The major steps include landmark and centerline feature extraction, indexing, and alignment for verification and refinement. The feature extraction algorithm is similar to the offline algorithm, but with an implementation optimized for speed: integer arithmetic, larger steps between centerline points on the blood vessels, and less precision in landmark location. The indexing step and the verification and refinement steps are described in detail in this section.

5.1   Indexing

Given a set of landmark locations extracted by the tracing algorithm, indexing is used to determine sets of possible correspondences between groups of these locations and groups of landmarks stored in the spatial map. The first step is to form landmark constellations using the same rules as for the diagnostic images. A quasi-invariant feature vector (QIFV)

8

is computed from each constellation. For each QIFV, $\mathbf{f}$, the indexing process itself is a search of the k-d tree to find the $m = 5$ stored vectors closest to $\mathbf{f}$. Each determines a set of correspondences between the constellation from the on-line image and from one of the diagnostic images stored in the spatial map.

Once the m nearest neighbors to a given on-line QIFV are found, they are ordered by increasing distance in feature space and stored individually. Unlike many object recognition techniques [3], the QIFV matches for different constellations are not clustered to generate combined hypotheses that could then be ranked-ordered. This is because nearly all of the landmarks detected during on-line tracing should correspond to landmarks detected off-line, and therefore a large fraction of matches should succeed. The match between off-line and on-line constellations generated from the first nearest neighbor for the first QIFV is tested and verified before all others. If this leads to a verified transformation, the whole spatial referencing procedure ends in success. Otherwise, the first nearest neighbor for the next QIFV is tested rather than the second nearest neighbor for the first. Only if the first nearest neighbor for all QIFVs fail are the second nearest neighbors considered. The reason for this is that the most likely cause of failure is an incorrect landmark [28].

5.2   Verification and Refinement Procedures

A match between a QIFV, $\mathbf{f}$, computed on-line and one stored in the k-d tree determines location correspondences which are used to estimate the four parameters of a similarity transformation. This transformation is a simple first step toward estimating the quadratic transformation between the entire on-line image and the spatial map. The transformation must be tested for accuracy (verification) and then refined (alignment) to increase its accuracy. These two steps are repeated as the transformation is increased in complexity from similarity to affine to quadratic, and is expanded to apply to the entire on-line image. Both the verification and refinement procedures depend on how accurately the current transformation aligns centerline trace points from the off-line image with trace points stored in the

spatial map. Centerline traces are used instead of landmarks for verification and alignment because they provide over an order of magnitude more constraints.

Each centerline trace location $\mathbf{x}_i$ from the on-line image is transformed to the spatial map based on the current transformation estimate. From the resulting location $\mathbf{x}'_i$ the closest centerline point $\mathbf{u}_i$ is found quickly using the pre-computed distance map. For verification, the median of the distances between all pairs of points $\mathbf{x}'_i$ and $\mathbf{u}_i$ is compared to an alignment accuracy threshold. For refinement, the pair of points $\mathbf{x}'_i$ and $\mathbf{u}_i$ becomes a (temporary) match. Using $\mathbf{u}_i$ and the location curve normal $\hat{\boldsymbol{\eta}}_i$ as a local, linear approximation to the centerline contour, the correspondence generated is effectively a point-to-line correspondence. These are used to refine the transformation by minimizing the robust error norm

$$E(\mathbf{p}) = \sum_i \rho\big([\mathbf{u}_i - \mathbf{M}(\mathbf{x}_i; \mathbf{p})] \cdot \hat{\boldsymbol{\eta}}_i \, / \, \sigma\big), \tag{2}$$

where $\mathbf{p}$ is the set of parameters determining the quadratic, affine, or similarity transformations (Section 2), $\rho(\cdot)$ is the subquadratic loss function of an M-estimator [14, 20, 30], and σ is a robustly estimated error standard deviation. Robustness is used to guard against spurious traces in the on-line image and missing traces in the spatial map. Typically, minimization of (2) uses iteratively reweighted least-squares [15], with only a few iterations. Convergence of the alternating matching and minimization procedures also takes only a few iterations.

5.3 On-Line Spatial Referencing Summary

As described above, the overall algorithm is greedy in taking the first correspondence generated from invariant indexing, and testing it immediately with the verification and refinement procedures. The mixture of steps in moving through verification and refinement, changing the transformation model, and expanding to cover the entire on-line image is as follows: (1) Local verification between the on-line image and the off-line image from which the matching constellation is taken. The similarity model is used and the transformation is applied to the

traces in a small region surrounding the constellation. (2) Finding the off-line image that is closest to the on-line image. The initial similarity transformation is used in combination with the offline transformations to find an initial estimate of the transformation between the on-line image and each off-line image. The off-line image that appears to have the closest viewpoint to the on-line image is chosen in order to minimize the distortion. (3) Initial refinement between the on-line image and the chosen off-line image. The affine model is used, initialized from the similarity transformation, and the region surrounding the constellation set is again used. (4) Verification of this "local" affine model. (5) Refinement of the affine transformation to apply to the entire on-line image. (6) Verification of the "global" affine transformation. (7) Refinement of the global quadratic transformation, initialized from the affine transformation. (8) Verification of the global quadratic transformation.

If the final verification succeeds, the transformation is accepted and spatial referencing is said to succeed. The thresholds on verification are chosen from plots of Receiver Operation Characteristic (ROC) [27] curves to minimize the number of false match acceptances [28]. If the verification fails at any stage, the next set of correspondences generated by invariant indexing is considered.

Overall, the algorithm is a realization of a robust iterative closest point algorithm [4, 12], specialized to aligning the retina. These specializations include the invariant indexing, the use of the spatial map, and the combined alignment and verification strategy described in the above outline.

6   Experimental Analysis

We have 18 sets of images of healthy retinas captured using a diagnostic fundus camera system at the Center for Sight, Albany, NY. Each set contains 19-22 images giving different views of the retina, for a total of 374 images. We do not yet have on-line sequences of images during surgery because implementation of the system to acquire such images is currently

i	1	2	3	4	5
λ_i	0.330	0.094	0.066	0.057	0.050
λ_i/λ_1	1.00	0.284	0.200	0.173	0.152

i	1	2	3	4	5	6
λ_i	1.59	1.41	0.246	0.217	0.204	0.186
λ_i/λ_1	1.00	0.888	0.155	0.137	0.128	0.117

Table 1: Principal component analysis of the triples invariants (top) and pairs invariants (bottom).

underway. The data set we do have is representative of the eventual on-line images because the imaging conditions will be similar.

Using the 18 data sets, we have developed a method of approximating the conditions under which spatial referencing must work. We refer to our test methodology on the 18 data sets as "leave-one-out". Suppose for a given data set there are n images, $I_1, \ldots, I_n$. One separate test is run for each image I_j. The image set $S_j = \{I_i | i \in [1 \ldots n], i \neq j\}$ is treated as the set of the diagnostic images from which the spatial map is built, and I_j is treated as the on-line image. The spatial referencing algorithm is then applied to I_j. Performance statistics are gathered by repeating this for all values of j and all data sets. On-line registration results are compared to the off-line, hand-validated registrations [11, 10] which we consider as a "gold standard". An example spatial referencing result appears in Figure 5.

6.1 Analyzing the Invariants

A preliminary test is to analyze the structure of the space of invariant indices. The distributions of the sets of pair and of triple QIFVs are analyzed using Principal Component Analysis (PCA). Angles are normalized to the range $[-1 \ldots 1)$ to make them comparable to location invariants (Section 3). Covariance matrices for all datasets are combined. Eigenvalues of these matrices are shown in Table 1.

The PCA of the pair invariants reveals two strong components and four weaker components. This is a result of the fact that the branching angle of the vasculature tends to be the same throughout the retina [18], so only the first angle is distinctive. On the other hand, the 3rd component is approximately equivalent (in ratio, at least) to the 5th component of the PCA of the triples QIFVs. This indicates that the effective dimension of the triples QIFVs is about two higher than that of the pairs. Remember, though, that there are about 4 times fewer pairs constellations, which counteracts this reduction in dimension.

6.2   Optimizing for Accuracy

Now we turn to statistical evaluation of the algorithm as a whole. The first evaluation uses a tracing algorithm implementation optimized for completeness and accuracy in the centerline locations and landmarks. This algorithm is always used for off-line construction of the spatial map, but it can also be used on-line to give an indication of ultimate effectiveness of the algorithm. In fact, with this tracing algorithm, the algorithm achieved *100% success*, which means that every "leave-one-out" image could be registered against the spatial map. The median registration error (compared to the gold standard) on 1024×1024 images was 0.94 pixels and the average error was 1.12. For the invariants from pairs, the error rates were 0.94 and 1.16, respectively.

6.3   Optimizing for Speed

The second version uses the implementation of tracing optimized for speed. This test is an important challenge because our goal is frame-rate performance. The images were also downsampled to 512×512. Such down-sampling is possible on the camera itself, so no loss in performance occurs. The overall success rate here is 96.5% for both triples and pairs. The median registration error is 1.05 pixels and the average error is 1.21, but the computation requires only 52 ms on average on a 900MHz Pentium III processor!

The drop in the success rate and the increase in error (especially in light of the reduction
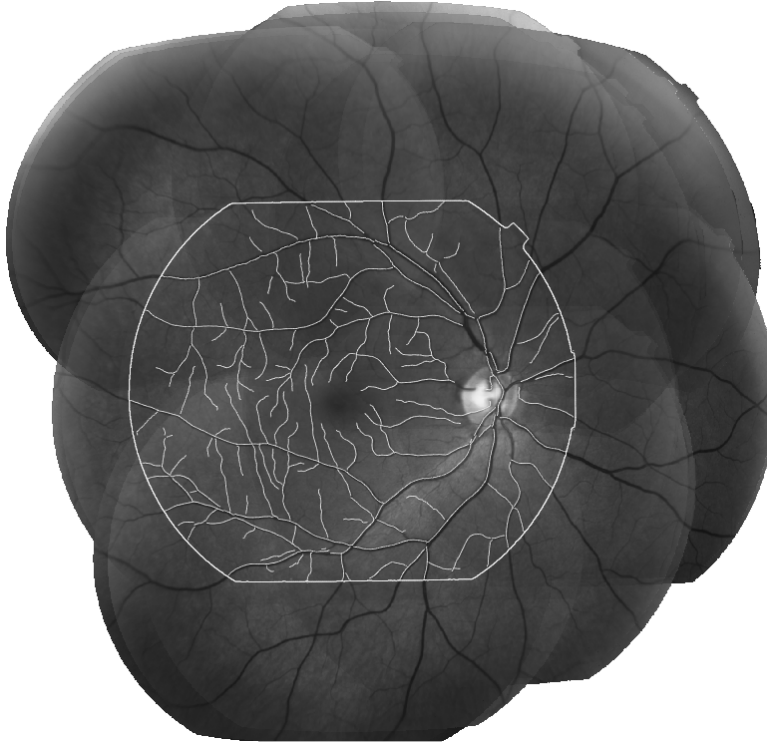
Figure 5: Example spatial referencing result showing the traces from the on-line image transformed onto a mosaic built from the off-line diagnostic images and transformations stored in the spatial map. This mosaic is centered on the diagnostic image chosen in step 2 of verification and refinement.

in resolution) are a direct result of using an implementation optimized for efficiency: the uncertainty of landmark and trace locations approximately doubles as compared to the off-line version. The effect of this can be seen by analyzing the images for which spatial referencing failed: in approximately 75% of the failures, no hypotheses made it past the first stage of verification. Most of these failures occurred for images on the periphery of the retina, where there are substantially fewer landmarks to begin with and therefore more sensitivity to errors or missing landmarks.

6.4   Discussion

These results are both promising and challenging. The 100% success rate with full-resolution images and off-line tracing shows the viability of the approach. The 52 millisecond computation time and 96.5% success rate on half-resolution images shows that we are near frame-rate

14

performance already. Indeed, combining the current implementation with a high-speed tracking algorithm would already provide a useful system for assisting in retinal laser surgery. The remaining challenges, of course, are improving the on-line success rate and the speed of our spatial referencing algorithm.

7   Summary and Conclusions

We have introduced an important new "spatial referencing" algorithm for high-speed registration of on-line retinal images against a spatial map of the retina, pre-computed from off-line, diagnostic images. This algorithm is fundamentally different from the tracking techniques that have been used in the past because it determines the absolute position of the on-line image without requiring knowledge of previous position estimates. This capability is crucial because unavoidable events — blinking, glare, head movements, saccadic eye movements — cause temporary loss of image quality and large changes in position between successive "good" images. The algorithm itself uses a combination of high-speed feature extraction, quasi-invariant indexing, greedy and aggressive testing of correspondences generated through indexing, and local-to-global verification and refinement of estimated transformations using iterative matching, robust estimation and increasing complexity transformations.

The transformation between the on-line image and the spatial map should provide the information required to determine the position of surgical tools, easily found in the on-line image, in the global, reference coordinate system defined by the spatial map. Many applications in the diagnosis and treatment of retinal diseases could be improved if this information were available. Our current focus is retinal laser surgery. Based on the results shown in this paper, our research is proceeding in several directions, with the goal of constructing a prototype treatment system. These directions include extensive experimentation with on-line image sequences, redesign of the vascular tracing algorithm to compute just sufficient information for spatial referencing, and redesign of ophthalmic imaging and treatment devices.

Acknowledgments

References

[1] S. F. Barrett, M. R. Jerath, H. G. Rylander, and A. J. Welch. Digital tracking and control of retinal images. *Optical Engineering*, 1(33):150–159, Jan. 1994.

[2] 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 Transactions on Biomedical Engineering*, 45(1):105–118, 1998.

[3] J. S. Beis and D. G. Lowe. Indexing without invariants in 3d object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10):1000–1015, 1999.

[4] P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[5] T. Binford and T. Levitt. Quasi-invariants: Theory and exploitation. In *Proceedings of the DARPA Image Understanding Workshop*, pages 819–829, 1993.

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

[7] 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*, 3(2):125–138, 1999.

[8] A. Can, C. Stewart, and B. Roysam. Robust hierarchical algorithm for constructing a mosaic from images of the curved human retina. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 286–292, 1999.

[9] A. Can, C. Stewart, B. Roysam, and H. Tanenbaum. A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: application to mosaicing the curved human retina. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 585–591, 2000.

[10] A. Can, C. Stewart, B. Roysam, and H. Tanenbaum. A feature-based algorithm for joint, linear estimation of high-order image-to-mosaic transformations: Mosaicing the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):412–419, 2002.

[11] A. Can, C. Stewart, B. Roysam, and H. Tanenbaum. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):347–364, 2002.

[12] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

[13] S. Fine. Observations following laser treatment for choroidal neovascularization. *Archives of Ophthalmology*, 106:1524–1525, 1988.

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

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

[16] D. Jacobs. Matching 3-d models to 2-d images. *International Journal 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] S. Kyriacos, F. Nekka, and L. Cartilier. Insights into the formation process of retinal vasculature. *Fractals*, 5(4):615–624, 1997.

[19] M. S. Markow, H. G. Rylander, and A. J. Welch. Real-time algorithm for retinal tracking. *IEEE Transactions on Biomedical Engineering*, 40(12):1269–1281, December 1993.

[20] P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim. Robust regression methods for computer vision: A review. *International Journal of Computer Vision*, 6:59–70, 1991.

[21] P. N. Monahan, K. A. Gitter, J. D. Eichler, and G. Cohen. Evaluation of persistence of subretinal neovascular membranes using digitized angiographic analysis. *Retina—The Journal of Retinal and Vitreous Diseases*, 13(3):196–201, 1993.

[22] P. N. Monahan, K. A. Gitter, J. D. Eichler, G. Cohen, and K. Schomaker. Use of digitized fluorescein angiogram system to evaluate laser treatment for subretinal neovascularization: Technique. *Retina—The Journal of Retinal and Vitreous Diseases*, 13(3):187–195, 1993.

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

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

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

[26] C. Rothwell, A. Zisserman, D. Forsyth, and J. Mundy. Planar object recognition using projective shape representation. *International Journal of Computer Vision*, 16(1):57–99, 1995.

[27] I. Selin. *Detection Theory.* Princeton University Press, 1965.

[28] H. Shen. *Indexing based frame-rate spatial referencing algorithms: application to laser retinal surgery.* PhD thesis, Rensselaer Polytechnic Institute, 2000.

[29] H. Shen, B. Roysam, C. Stewart, J. Turner, and H. Tanenbaum. Optimal scheduling of tracing computations for real-time vascular landmark extraction from retinal fundus images. *IEEE Transactions on Information Technology in Biomedicine*, 5(1):77–91, Mar 2001.

[30] C. V. Stewart. Robust parameter estimation in computer vision. *SIAM Reviews*, 41(3), September 1999.

[31] C.-L. Tsai, C. Stewart, B. Roysam, and H. Tanenbaum. Repeatable vascular landmark extraction from retinal fundus images using local vascular traces. *IEEE Transactions on Information Technology in Biomedicine*, submitted 2001.

[32] 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.

[33] A. Zisserman, D. Forsyth, J. Mundy, C. Rothwell, J. Liu, and N. Pillow. 3d object recognition using invariance. *Artificial Intelligence*, 78(1-2):239–288, 1995.