The Dual Bootstrap Iterative Closest Point Algorithm with Application to Retinal Image Registration[*]

Charles V. Stewart[1]   Chia-Ling Tsai[1]   Badrinath Roysam[2]

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

[2]Dept. of Electrical, Computer,
and
Systems Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180–3590
roysab@rpi.edu

May 28, 2003

Abstract

Motivated by the problem of retinal image registration, this paper introduces and analyzes a new registration algorithm called Dual-Bootstrap Iterative Closest Point (ICP). The approach is to start from one or more initial, low-order estimates that are only accurate in small image regions, called bootstrap regions. In each bootstrap region, the algorithm iteratively (a) refines the transformation estimate using constraints only from within the bootstrap region, (b) expands the bootstrap region, and (c) tests to see if a higher-order transformation model can be used, stopping when the region expands to cover the overlap between images. Steps (b) and (c), the bootstrap steps, are governed by the covariance matrix of the estimated transformation. Estimation refinement (a) uses a novel robust version of the Iterative Closest Point (ICP) algorithm. In registering retinal image pairs, Dual-Bootstrap ICP is initialized by automatically matching individual vascular landmarks, and it aligns images based on detected blood vessel centerlines. The resulting quadratic transformations are accurate to less than a pixel. On tests involving approximately 6000 image pairs, it successfully registered 99.5% of the pairs containing at least one common landmark, and 100% of the pairs containing at least one common landmark and at least 35% image overlap.

Index terms: Registration, iterative closest point, robust estimation, retinal imaging, medical imaging

---

[*]A preliminary version of this paper appeared in [80]

1 Introduction

This paper introduces a new algorithm called Dual-Bootstrap Iterative Closest Point (ICP) and uses it to solve the retinal image registration problem. The new algorithm is based on a new approach to registration, which we call the "dual-bootstrap." We start with the retinal image registration problem and use this to motivate the algorithm.

1.1 The Retinal Image Registration Problem

Images of the retina are used to diagnose and monitor the progress of a variety of diseases, including such leading causes of blindness as diabetic retinopathy, age-related macular degeneration, and glaucoma. These images (Figure 1) are usually acquired using a fundus camera looking through the lens of the eye, and are taken in both the visible spectrum and, using fluorescein and indocyanine green angiography, in the near infrared [6, 65]. A variety of imaging protocols are used [24] to produce images showing various parts of the retina. Angiography sequences reveal the flow of blood through the retina and are therefore used to highlight blockages and weak, leaking vessels.

Retinal image registration has a variety of applications (Figure 2). Registering a set of images taken during a single session with a patient can be used to form a single, composite (mosaic) view of the entire retina [14, 23, 83]. Multimodal registration can reveal the relationship between events seen on the surface of the retina and the blood flow shown in the angiography. Registering images taken weeks, months or years apart can be used to reveal changes in the retina at the level of small regions and individual blood vessels.

Retinal image registration is challenging. The images are projections of a curved surface taken from a wide range of viewpoints [24] using an uncalibrated camera. The non-vascular surface of the retina is homogeneous in healthy retinas, and exhibits a variety of pathologies in unhealthy retinas. Unfortunately (for the purposes of registration), these pathologies can appear and disappear over time (Figure 3), making them poor choices for longitudinal registration. Only the vasculature covers the entire retina and is relatively stable over time.

Thus, it appears that a solution to the retinal image registration problem requires an approach driven by the vascular structure. This can include both the vessels themselves and their branching

1

and cross-over points (Figure 3). Choosing to use the vasculature does not make the problem easy, however. There are many vessels and many of these locally appear similar to each other. The effects of disease and poor image quality can obscure the vasculature. Moreover, in different stages of an angiography sequence, different blood vessels can be bright, while others are dark. Finally, the range of viewpoints dictated by some imaging protocols implies the need to register image pairs having small amounts of overlap. Together, these observations imply that (1) initialization is important, (2) minimization will require avoiding local minima caused by misalignments between vessels, and (3) minimization must also be robust to missing structures. These problems are common to many registration problems. They are addressed here in a novel way using the Dual-Bootstrap ICP algorithm.

1.2   Dual-Bootstrap ICP

Consider the images shown in Figure 3. Automatic detection of vascular landmarks (branching and cross-over points in the vessel) produces just one landmark in common between these two images. Aligning the images based on the common landmark and then running an iterative minimization algorithm (a robust version of ICP [7, 17, 87]) results in substantially misaligned images (Figure 4). The initialization is too poor globally (image-wide) for the minimization to correctly align the complex vascular structure. Intuitively, this failure is unsatisfying, however, because the alignment in the small initial region (the white rectangle in Figure 4) is reasonably accurate. What is needed is a way to expand this locally-accurate initial alignment into a globally-accurate final alignment. Dual-Bootstrap ICP algorithm does exactly this.

The Dual-Bootstrap ICP algorithm starts from an initial transformation estimate that is only accurate in a small region, R, (the "bootstrap region") of the mapped image, and expands it into a globally-accurate final transformation estimate. This expansion iteratively refines and extends the transformation. This process, illustrated for the above retina image registration example in Figure 5, is accomplished by repeating three major steps:

Estimating the transformation: The transformation is estimated only in the bootstrap region. Dual-Bootstrap ICP uses a robust form of ICP, but other transformation estimation algorithms could be used, especially in different application domains.

Region bootstrapping: Based on the uncertainty in the transformation estimate, the bootstrap region is expanded. This is governed by the transformation parameter estimate covariance matrix. Stable, accurate estimates cause rapid growth, while unstable, inaccurate estimates cause slow growth.

Model bootstrapping: Rather than using a single, fixed transformation model, different models are used as the bootstrap region expands, starting from a simple model for the initial bootstrap region and gradually evolving to a higher-order model as the bootstrap region grows to cover the entire image overlap region. Model selection techniques [11, 76] are used to automatically select the transformation model for each bootstrap region. The set of transformations used in retinal image registration is shown in Table 1.

Thus, the term dual-bootstrap refers to simultaneous growth in the bootstrap region and the transformation model order. Initialization of the bootstrap region can be accomplished in many ways, including manually specification of common structures by a clinician. In the retina application, where the entire registration procedure is fully-automatic, landmarks are matched based on invariant signatures, and the initial bootstrap region is formed around the landmarks. The dual-bootstrap procedure is run on several such initial bootstrap regions. If the procedure converges to a sufficiently-accurate image-wide transformation for one of these regions, the two images are considered to be accurately aligned and registration succeeds.

2 Background

2.1 Approaches to Registration

Registration is a fundamental problem in automatic image analysis [10]. A number of useful surveys of registration within the medical imaging domain exist [31, 41, 47]. Many medical image registration techniques address the problem of accurate alignment of intra- and inter-modality images given reasonable starting estimates [31, 45, 82]. Other research in medical image registration focuses on the deformations necessary to align images, taking cues from physics-based models [47, 49]. In many applications such as retinal image registration the most important issues

are initialization, convergence, and robustness to missing and misaligned structures; handling substantial deformations is less important.

Initialization can be addressed in a variety of ways, including image-wide measurements [36, 40], multiresolution [5, 27, 66, 67], indexing and initial matching of distinctive features or sets of features [18, 21, 44, 69, 75], and minimal-subset (of possible correspondences) random-sampling techniques [28, 34, 64, 78, 79, 88].

The major distinction in minimization methods is between intensity-based and feature-based approaches. Intensity-based approaches generally optimize an objective function based on comparison of intensities or intensity gradients [59], or based on measures such as mutual information [31, 45, 82]. Feature-based techniques align images based on correspondences between automatically detected features [15]. In the retina application, the stable and prominent vascular structure should drive the minimization. This is done using a feature-based method here, but other methods are certainly possible. Of particular interest is the idea of aligning vascular features of one image with the intensity structure of another [2, 3]. We can think of this as a partial feature-based approach.

Dual-Bootstrap ICP uses the ICP algorithm, which was invented almost simultaneously in the early 1990's by several different research groups [7, 16, 17, 51, 87] and has been used in many applications since then [25, 42, 51, 56]. ICP is based on point features, where the "points" may be raw measurements such as (x, y, z) values from range images, intensity points in three-dimensional medical images [27, 30], and edge elements, corners and interest points [68] that locally summarize the geometric structure of the images. The idea of ICP is to alternate steps of (1) closest point matching based on a transformation estimate and (2) transformation estimation based on the resulting matches until convergence to a local minimum. Clearly, good initialization is an important precursor to ICP.

2.2 Retinal Image Registration

Retinal image registration algorithms in the literature differ based on the the transformation models, the image primitives, and the search techniques.

The transformation models applied in retinal image registration include translation [4, 50, 84],

Euclidean [43, 57], similarity [20, 62], affine [33, 39, 48, 50, 54, 60, 62, 85], bilinear [48, 54] and quadratic [15, 26, 35, 46, 70]. [15] derives the quadratic transformation model based on a weak perspective camera model and a quadratic approximation to the retinal surface. Table 1 summarizes the accuracies of several retinal image transformation models, showing that for 1024×1024 images and a wide range of overlaps, the quadratic transformation is needed for accurate alignment. For lower resolutions and a high fraction of image overlap, lower-order transformations can be used.

As in the general registration problem, transformation parameter estimation techniques for retinal image registration can be classified into intensity-based [20, 48, 54, 58, 62], and feature-based [4, 15, 26, 33, 35, 39, 43, 46, 50, 57, 60, 70, 84, 85]. Intensity-based approaches have used intensity differences on segmented images [48, 54], and cross-correlation[20, 58] and mutual information measures[62] on the original images. These algorithms have been embedded in non-gradient, global search techniques such as simulated annealing or genetic algorithms because of the large number of local minimum in the search space [48, 54]. The resulting techniques require many function evaluations (up to 10,000) for affine transformations. This problem is substantially worse for quadratic transformations. Intensity-based approaches can also suffer from background changes due to pathologies or artifacts.

Feature-based methods usually minimize an objective function of the distances between vascular landmark correspondences. To narrow down the search space to the correct domain of convergence, [26, 33, 46, 84] initialize the transformation using "matched" landmarks, such as the optic disk and the fovea in [26, 33, 84], manually specified points [46], and triples of landmarks assuming the images are initially close [39]. [35, 57, 85] estimate the most plausible transformation from a pool of possible landmark matches using techniques such as the Hough transform [57, 85], expectation-maximization [35], and voting [4]. [43, 60] guess several different initial transforms from seeds or Fourier measures and then refine based on alignment of vessels using non-gradient-based minimization techniques. The form of these constraints is roughly similar to those in the Dual-Bootstrap ICP without the dual-bootstrap procedure for avoiding local minima. [50] searches for the corresponding areas based on template matching on edges. Having the initial transformation, most algorithms iteratively remove the outliers based on a

distance constraint and refine the transformation using techniques such as least-squares minimization. Our group has recently published a landmark-based retinal registration algorithm that aligns automatically-detected landmarks using a hierarchical estimation technique [13, 15]. Even more recently, we have described a real-time algorithm designed to align images taken during surgery with images forming a spatial map of the retina [69, 70]. This algorithm uses invariant indexing and alignment of blood vessels for refinement, but does not use dual-bootstrapping. In all landmark-based algorithms, the major problem is finding and matching a set of consistent landmarks to sufficiently constrain the transformation needed for accurate registration. This is particularly a problem for non-linear transformations, low image overlap, poor quality images, and longitudinal changes.

In summary, while a wide variety of techniques has been proposed for the retinal image registration problem, none of them yet address all the issues outlined in Section 1.1.

2.3   Multi-Image Registration

Many retinal imaging protocols require acquisition of multiple images [24]. These must be aligned in a manner that is mutually consistent. Hence, any algorithm that registers pairs of retinal images ("pairwise registration") is not a complete retinal image registration system [4, 14, 46, 60]. A second algorithm is needed to ensure mutual consistency of all aligned images [14]. Like many other multi-image alignment techniques [66, 71], the algorithm in [14] uses constraints generated from pairwise registration to initialize an adjustment procedure that results in transformations between all pairs of images. Image pairs that fail pairwise registration may even be aligned accurately if the final alignment of the images involved is sufficiently well-constrained by other pairwise registration results. We will exploit this fact to aid in the experimental evaluation of the Dual-Bootstrap ICP algorithm. The ultimate success of multi-image registration still depends on pairwise registration, however, and when there are only two images to be aligned, multi-image registration plays no role.

2.4 Retinal Vascular Feature Extraction

Many techniques have been proposed in the research literature for vascular feature extraction [12, 38, 60, 72, 74, 86]. In the implementation of Dual-Bootstrap ICP, we employ an algorithm that extracts elongated structures using two-sided boundary following [1, 12, 29, 81], which we have termed "tracing". This algorithm is fast (less than a second), fully-automatic, and adapts automatically to the contrast in each image region. The features extracted are trace point locations along the vessel centerlines, akin to medial axis points. Each centerline (trace) point is described by a location, tangent direction, and width. Branching and cross-over points ("landmarks") of the vessels are extracted as well. Each landmark is described by its image coordinates and the orientations and widths of the vessels that meet to form it. Other feature extraction techniques could be used in the Dual-Bootstrap ICP algorithm in place of our own. Moreover, as illustrated in Figure 5, much of the registration algorithm design is intended to recover from poor quality features.

3 The Dual-Bootstrap ICP Algorithm

The organization of the algorithm description has a two-fold purpose. The first is to show how Dual-Bootstrap ICP is used to register retinal images. The second is to present the core dual-bootstrap technique as a procedure that might be used in a variety of other applications. Following an introduction to the notation and a procedural outline, Sections 3.3, 3.7, and 3.8 are specific to the retinal application, while the description of the core procedure in Sections 3.4-3.6 is more general.

3.1 Notation

The following notation will be used:

- Let I_1 be the image being mapped (the "moving" image) and let I_2 be the fixed image (the "target" image) that I_1 is being mapped onto.

- The centerline trace point and landmark feature sets will be denoted $\mathcal{P}_i$ and $\mathcal{L}_i$, respectively for $i = 1, 2$.

7

- R_t will denote the bootstrap region at iteration t of the dual-bootstrap procedure. Bootstrap regions are defined in the coordinate system of image I_1. The initial bootstrap region is R_1.

- $M(\boldsymbol{\theta}; \mathbf{p})$ is a transformation function mapping image location (or feature vector) $\mathbf{p}$ from the coordinate system of I_1 onto the coordinate system of I_2. Here $\boldsymbol{\theta}$ is the set of transformation parameters to be estimated.

- Let $\mathcal{M}$ be a sequence of such transformation functions or "models". The model set may or may not form a nested hierarchy. In the retina application, $\mathcal{M}$ is {*similarity, affine, reduced-quadratic, quadratic*} (Table 1).

- M_t will denote the model selected in bootstrap region, R_t, during iteration t. $\hat{\boldsymbol{\theta}}_t$ is the estimated vector of parameters instantiating the model M_t.

- Σ_t is the covariance matrix of the estimated parameter vector $\hat{\boldsymbol{\theta}}_t$.

- $E(R_t, M_t, \boldsymbol{\theta}_t)$ is the registration objective function that measures the transformation error between I_1 and I_2 in region R_t using model M_t and parameter vector $\boldsymbol{\theta}_t$.

3.2   Procedure

The following outlines the entire Dual-Bootstrap ICP procedure:

1. Pre-computation: Apply feature extraction to each image, I_1 and I_2, to produce the feature sets $\mathcal{P}_i$ and $\mathcal{L}_i$, $i = 1, 2$. Match landmarks between $\mathcal{L}_1$ and $\mathcal{L}_2$ to produce initial correspondences.

2. For each initial correspondence:

   (a) Initialize the model M_1 to be the lowest order model, and compute an initial transformation estimate. Establish the initial bootstrap region R_1.

   (b) $t = 1$

   (c) (Dual-Bootstrap procedure) While the estimate has not converged:

8

i. Estimate parameter vector $\hat{\boldsymbol{\theta}}_t$ by minimizing $E(R_t, M_t, \boldsymbol{\theta}_t)$ using a robust ICP algorithm. Calculate the covariance matrix $\boldsymbol{\Sigma}_t$ of the estimate $\hat{\boldsymbol{\theta}}_t$.

ii. Bootstrap the model: Apply a model selection technique to choose the new model M_{t+1}. If the model selection technique chooses a different model — that is, $M_t \neq M_{t+1}$ — then $\hat{\boldsymbol{\theta}}_t$ and $\boldsymbol{\Sigma}_t$ must be replaced by the estimate and covariance computed for M_{t+1} during model selection.

iii. Bootstrap the region: Use the covariance matrix, $\boldsymbol{\Sigma}_t$, and the new model M_{t+1} to expand the region based on the "transfer error".

iv. Check for convergence.

v. $t = t + 1$

(d) If the converged transformation estimate $\hat{\boldsymbol{\theta}}$ is sufficiently accurate, terminate with a successful registration.

3. When no more initial estimates are available, terminate and indicate that the images can not be registered.

3.3   Initialization: Invariants and Initial Estimates

A brief description of the invariant signature computation and matching (1) and initial estimation (2a) steps is as follows. Matches between two landmarks, one in each image, or between pairs of landmarks in each image are generated by computing and comparing invariants [8, 55]. Invariants for a single landmark are blood vessel width ratios and blood vessel orientations (Figure 6), giving a five-component invariant signature vector.[1] The invariant signature of a set of two landmarks is a six-component vector. The line segment drawn between the two landmarks forms an axis, and the orientation of each of three landmark angles is computed with respect to this axis, giving the six components.[2] The invariant signature vectors for one- and two-landmark sets are computed separately for each image, I_1 and I_2, and then matched (each set separately). The closest match is found for each signature vector. Additional matches are determined when the Mahalanobis distance between signature vectors is within a 95% confidence chi-squared uncertainty bound.

---

[1]These are invariant to translation and scaling, making them quasi-invariant to higher order models.
[2]These are similarity invariants.

Each signature match produces a set of one or two landmark correspondences, depending on whether single landmark or landmark pair invariants were matched. These sets are ordered for testing by chi-squared confidence levels.

For each set of landmark correspondences chosen (step 2) to initialize the dual-bootstrapping procedure, parameters of a similarity transformation are estimated (step 2a) from the landmark locations and the orientations and widths of the vessels that meet to form the landmarks. For single-landmark correspondences, the initial bootstrap region, R_1, is an axis-aligned square centered on the landmark in I_1. The width of the square is $\alpha = 10$ times the width of the widest vessel meeting to form the landmark. For pairs of landmark correspondences, R_1 is the smallest axis-aligned rectangle containing the squares formed around both landmarks in I_1.

3.4 Estimating the Transformation Parameters

The first step in the dual-bootstrap loop (2c) is estimating the transformation within the bootstrap region R_t. Since registration is based on blood vessel trace points, the objective function to be minimized may be written as:

$$E(R_t, M_t, \boldsymbol{\theta}_t) = \sum_{\mathbf{p} \in \mathcal{P} \cap R_t} \rho(d(M_t(\boldsymbol{\theta}_t; \mathbf{p}), I_2)/\hat{\sigma}). \tag{1}$$

Here, $d(M_t(\boldsymbol{\theta}_t; \mathbf{p}), I_2)$ is the distance between the transformation of $\mathbf{p}$ and the closest point (trace) in I_2, $\rho(u)$ is a robust loss function [32, 73], monotonically non-decreasing as a function of $|u|$, and $\hat{\sigma}$ is the error scale, which is the (robust) standard deviation of the error distances.

3.4.1 Robust ICP

In ICP the objective function is minimized by alternating steps of matching and parameter estimation. In matching, trace points $\mathbf{p} \in \mathcal{P}_1$ that are also in R_t are mapped into I_2 based on the current transformation. This forms point $\mathbf{p}' = M_t(\hat{\boldsymbol{\theta}}_t; \mathbf{p})$. The closest image I_2 trace point $\mathbf{q} \in \mathcal{P}_2$ to $\mathbf{p}'$ is found (Figure 7). Digital distance maps are used to accelerate matching [9, 22]. The distance metric is

$$d(M_t(\boldsymbol{\theta}_t; \mathbf{p}), \mathbf{q}) = |(M_t(\boldsymbol{\theta}_t; \mathbf{p}) - \mathbf{q})^T \boldsymbol{\eta}_q|, \tag{2}$$

10

where $\boldsymbol{\eta}_q$ is the normal to the trace at $\mathbf{q}$. This gives a "point-to-line" distance, which is appropriate for matching points on vessel contours. Each resulting correspondence, $(\mathbf{p}, \mathbf{a})$ is placed in the correspondence set $\mathcal{C}_t$ for this iteration of ICP (not the iteration of the dual-bootstrap loop).

The new transformation parameter estimate (for this iteration of ICP) $\hat{\boldsymbol{\theta}}_t$ is calculated based on rewriting (1) using the correspondences and the distance metric (2):

$$E(\mathcal{C}_t, \boldsymbol{\theta}_t) = \sum_{(\mathbf{p},\mathbf{q}) \in \mathcal{C}_t} \rho([M_t(\boldsymbol{\theta}_t; \mathbf{p}) - \mathbf{q}]^T \boldsymbol{\eta}_q)/\hat{\sigma}). \tag{3}$$

The equation is minimized using iteratively-reweighted least-squares (IRLS) [37], with weight function $w(u) = \rho'(u)/u$. The minimization process alternates weight recalculation using a fixed parameter estimate with weighted least-squares estimation of the parameters.

The choice of loss functions is motivated by looking at several different loss functions and their associated weight functions illustrated in Figure 8. The least-squares loss function has a constant weight, the Cauchy weight function descends and asymptotes at 0, while the Beaton-Tukey biweight function [37] has a hard limit beyond which the weight is 0. This limit is set to about $4\hat{\sigma}$ in the statistics literature. This is important for rejecting errors due to mismatches, and therefore it is the one we choose (other loss functions sharing this hard-limit property could also be used). In detail, the weight function is

$$w(u) = \begin{cases} [1 - \left(\frac{u}{4}\right)^2]^2 & |u| \le 4 \\ 0 & |u| > 4, \end{cases} \tag{4}$$

where $u = d(M_t(\boldsymbol{\theta}_t; \mathbf{p}), \mathbf{q})/\hat{\sigma}$.

3.4.2 Robust Error Scale Estimation

Estimation of error scale, σ, is done for each set of correspondences, $\mathcal{C}_t$, at the start of reweighted least-squares. We use a technique called MUSE that automatically adjusts its estimate by determining the fraction of (approximately) correct matches [52, 53]. This is important because sometimes more than 50% of the feature points in R_t are mismatched. (An example of this occurs during the registration process shown in Figure 5 when R_t covers about half the overlap between

11

images.) Let $r_j = |d(M_t(\hat{\boldsymbol{\theta}}_t; \mathbf{p}_j), \mathbf{q}_j)|$ be the absolute error estimate for the jth correspondence using the current estimate $\hat{\boldsymbol{\theta}}_t$ of the transformation parameters. Let $r_{1:N}, r_{2:N}, \ldots, r_{N:N}$ be a rearrangement of these values into non-decreasing order. Then for any k, $r_{1:N}, \ldots, r_{k:N}$ are the k smallest errors. A scale estimate may be generated from $r_{1:N}, \ldots, r_{k:N}$ as

$$\sigma_k^2 = \frac{\sum_{j=1}^{k} r_{j:N}^2}{C(k, N)},$$

where $C(k, N)$ is a computed correction factor. This factor makes σ_k^2 an unbiased estimate of the variance of a normal distribution using only the first k out of N errors. The intuition behind MUSE is seen by considering the effect of outliers on σ_k^2. When k is large enough to start to include outliers (errors from incorrect matches), values of σ_k^2 start to increase substantially. When k is small enough to include only inliers, σ_k^2 is small and approximately constant. Thus, we can simply evaluate σ_k^2 for a range of values of k (e.g. $0.35N, 0.40N, \ldots 0.95N$), and choose the smallest σ_k^2. To avoid values of k that are too small, we take the minimum variance value of σ_k^2, not just the smallest σ_k^2. Details are in [52, Chapter 3].

3.4.3   Covariance Matrix

The final step in transformation parameter estimation, after the inner robust ICP loop is finished, is computing the covariance matrix of the estimated parameters. We approximate this from the inverse Hessian — the matrix of second partial derivatives — of the objective function evaluated at the current estimate [61, Ch. 15]. Intuitively, the Hessian measures the steepness of the minimum; hence, the variance of the parameters is low for steep minima. In equation form, the covariance matrix is

$$\boldsymbol{\Sigma}_t = \mathbf{H}^{-1}(E(R_t, M_t, \hat{\boldsymbol{\theta}}_t)). \tag{5}$$

where $E()$ is defined in (1). For a fixed set of correspondences, $E()$ simplifies to the expression in (3), which is used here. This is exact for quadratic ρ functions, and approximate otherwise. Finally, there is no need to multiply $\mathbf{H}^{-1}$ by $\hat{\sigma}^2$, because $\hat{\sigma}^2$ is already used to normalize the objective function.

3.5 Bootstrapping the Model

Small bootstrap regions R_t usually contain insufficient constraints to reliably estimate the parameters of the quadratic transformation (see Table 1). Moreover, non-linear distortions are generally small in these regions. For large bootstrap regions the non-linear distortions are significant, and the quadratic transformation is well-constrained. One possible solution to this would be to introduce stabilizing multipliers on the non-linear parameters of the quadratic model for small image regions and use just a single model. This would introduce one or more tuning constants, however, and would not help with under-constrained linear terms.

A second approach, and the one we adopt, is to use a set of transformations models, each successively more accurate and involving more parameters, and automatically switch between models as R_t expands. (See [77] for a similar approach in the context of multi-camera imaging geometry.) Table 1 shows the models used in retinal image registration, and illustrates the successive accuracy of each model. Changing the model order must be done carefully, however. Switching to higher order models too soon causes the estimate to be distorted by noise. Switching too late causes modeling error to increase, which increases the alignment error. This can cause misalignments on the boundaries of R_t, sometimes leading registration into the wrong domain of convergence. To select the correct model for each bootstrap region, statistical model selection techniques are applied.

Statistical model selection is a well-studied problem [11, 76], and experimental analysis shows that several techniques have approximately equivalent performance [11]. All techniques choose the model that optimizes a trade-off between the fitting accuracy of high-order models and the stability of low-order models. Stability is generally measured in terms of the covariance matrix of the parameters. (For a fixed data set, covariance matrices of higher-order models generally have smaller eigenvalues than covariance matrices of lower-order models and therefore should be viewed as less stable.) The current Dual-Bootstrap ICP model selection criteria is adapted from [11]. The criteria depends on the following expression:

$$\frac{d}{2} \log 2\pi - \sum_i w_i r_i^2 + \log \det(\Sigma), \tag{6}$$

13

where d is the degrees of freedom in the model, $\sum_i w_i r_i^2$ is the sum of the robustly-weighted alignment errors (based on the estimate $\hat{\boldsymbol{\theta}}$), and $\det(\Sigma)$ is the determinant of the parameter estimate covariance matrix. The t subscripts have been dropped in this equation.

Equation 6 is evaluated for the current model M_t and for other candidate models from the set $\mathcal{M}$. For each other model M, the objective function $E(R_t, M, \boldsymbol{\theta})$ must be minimized to produce the weighted alignment errors and the covariance matrix. For simplicity, we use the correspondence set $\mathcal{C}_t$ from the minimization of $E(R_t, M_t, \boldsymbol{\theta}_t)$ to reach an approximate minimum for each M rapidly. Also, for simplicity, we can just evaluate M_t and the next more complex model in $\mathcal{M}$, especially since region growth is monotonic. Overall, the model that results in the greatest value of (6) is chosen at M_{t+1}.

3.6 Bootstrapping the Region

Region expansion should be controlled by the uncertainty in the transformation estimate: uncertain transformations lead to uncertain mapping of feature points and therefore uncertain matching. (Errors in matching resulting from errors in the transformation are exactly what the dual-bootstrap procedure is designed to avoid.) This requires deriving the uncertainty in the feature point mapping from the covariance of the transformation parameters, and then developing region expansion equations from this.

The uncertainty in mapping of point locations is computed from the covariance of the transformation parameter estimate using the "transfer error" [34, Ch. 4]. Specifically, let $\mathbf{p}_i$ be a location in I_1, with coordinates computed with respect to the center of R_t. Let $\mathbf{p}_i' = M_t(\hat{\boldsymbol{\theta}}_t; \mathbf{p}_i)$ be the mapping of $\mathbf{p}_i$ onto image I_2 based on the current estimate. This mapped point is a random variable because of its dependence on the transformation.[3] Its covariance matrix, denoted by $\boldsymbol{\Sigma}_{\mathbf{p}_i'}$, can be approximated from $\boldsymbol{\Sigma}_t$ and the Jacobian of the transformation evaluated at $\mathbf{p}_i$:

$$\boldsymbol{\Sigma}_{\mathbf{p}_i'} = \left(\frac{\partial M(\hat{\boldsymbol{\theta}}_t; \mathbf{p})}{\partial \mathbf{p}}(\mathbf{p}_i) \right) \boldsymbol{\Sigma}_t \left(\frac{\partial M(\hat{\boldsymbol{\theta}}_t; \mathbf{p})}{\partial \mathbf{p}}(\mathbf{p}_i) \right)^T. \tag{7}$$

For two-dimensional retinal images this is a 2×2 positive definite matrix whose terms increase

---

[3]For simplicity, we do not treat $\mathbf{p}_i$ as a random variable here: variability in feature location is much less than variability in the transformation until the estimation converges.

with increasing distance from the center of the region.

Several region growth techniques could be developed using the transfer error. For example, we could define R_{t+1} to be the set of pixel locations in I_1 for which the trace of the transfer error covariance matrix is less than a threshold. This would require computing and mapping the covariance at each pixel. The measure we adopt here is simpler and faster to compute. Since the bootstrap regions are axis-aligned rectangles, we grow each side of R_t outward independently. This requires the computation of only four transfer error covariance matrices per iteration t.

In particular, growing any one of the four sides of R_t involves pushing $\mathbf{p}_i$, which we now define to be the center of the side, out along the side's normal direction, $\boldsymbol{\eta}_i$ (see Figure 9). In order to exploit the transfer error covariance, which is defined in the coordinate system of I_2, outward normal vector $\boldsymbol{\eta}_i$ must be mapped into I_2 and rescaled to produce the mapped unit normal, $\boldsymbol{\eta}_i'$. Projecting the transfer error covariance on this (outward) direction produces the scalar $\boldsymbol{\eta}_i'^T \boldsymbol{\Sigma}_{\mathbf{p}_i'} \boldsymbol{\eta}_i'$. Using this, we define the outward growth of the side in terms of the shift in $\mathbf{p}_i$ as:

$$\Delta \mathbf{p}_i = \beta \frac{(\mathbf{p}_i^T \boldsymbol{\eta}_i)}{\max(1, \boldsymbol{\eta}_i'^T \boldsymbol{\Sigma}_{\mathbf{p}_i'} \boldsymbol{\eta}_i')} \boldsymbol{\eta}_i. \tag{8}$$

This growth is proportional to the current distance $(\mathbf{p}_i^T \boldsymbol{\eta}_i)$ of the side $\mathbf{p}_i$ lies on from the center of R_t, and the growth is inversely proportional to the transfer error in the normal direction. The lower bound of 1 in the denominator prevents growth from becoming too fast. The center of each side of R_t is expanded outward independently using Equation 8, and the new region R_{t+1} is the axis-aligned rectangle formed by the resulting four points. Parameter β tunes the growth rate. A value of $\beta = \sqrt{2} - 1$ ensures that the area of a two-dimensional region at most doubles in each iteration.

3.7   Convergence and Termination

There are two levels of convergence and termination. First, the dual-bootstrap procedure for a single initial estimate (step 2c) ends when region R_t covers the entire apparent overlap region between images I_1 and I_2. Several heuristics are used to help identify incorrect alignments early in the dual-bootstrap procedure, and thereby accelerate the algorithm: region growth that is too

slow, alignment error (defined below) that is too large, or a condition number of the parameter estimate covariance matrix that is too low. If any of these occurs the procedure is terminated and a new initial estimate is tried. Thresholds are set conservatively to avoid eliminating correct estimates.

The second level is termination of the overall registration procedure. This occurs when (1) the alignment error of the final transformation estimate $\hat{\boldsymbol{\theta}}_t$ is sufficiently low and (2) the covariance matrix of the quadratic transformation is well-conditioned. Of these, the alignment error is by far the most important. The error, called the "Centerline Error Measure" or CEM, is defined as the median of the distances $d(M_t(\hat{\boldsymbol{\theta}}_t; \mathbf{p}), \mathbf{q})$ over all correspondences $(\mathbf{p}, \mathbf{q})$ in the final transformation set, $\mathcal{C}_t$. We use the acceptance threshold of 1.5 pixels on the CEM, as established empirically for a previous algorithm by Can [15].

3.8   Parameters

The entire Dual-Bootstrap ICP algorithm is largely parameter-free. Of course, there are statistical parameters such as the 95% confidence threshold on Chi-square distance measures, and the 4σ parameter in the Beaton-Tukey biweight, but these are not considered adjustable. The only parameters that could be considered adjustable are

- α — the initial region multiplier (Section 3.3),

- β — the region growth rate parameter (Section 3.6), and

- the CEM threshold (Section 3.7).

The first two were set at 10 and $\sqrt{2} - 1$ early in our studies and have not been changed, except to do a retrospective sensitivity analysis (see below). The CEM threshold of 1.5 taken from [15] was never varied.

4   Experimental Analysis

This section presents the results of a large number of experiments using Dual-Bootstrap ICP in retinal image registration. The presentation illustrates the nearly flawless performance of the

16

algorithm and then illustrates the importance of each major component of the dual-bootstrapping process.

4.1   Data and Analysis

The performance of the algorithm is demonstrated on two groups of image data sets. One contains images from 18 different healthy eyes, with 15-25 images in each set. These images, which were taken with a Topcon digital fundus camera, show a wide range of views of each retinal fundus, and some pairs of images of the same eye have no overlap whatsoever. This dataset was the basis for experiments reported with our recent, landmark-based algorithm [15]. The second data set contains images from 40 different eyes with various pathologies, yielding 300 image pairs. Some of these pairs were taken at the same time, while others were taken with time differences of months or even years. Some of these images are from a digital camera, while others are scanned color slides. All images have a resolution of 1024×1024 pixels. Results are presented for the two data sets separately because the second, "pathology" set is more challenging, but much smaller. Figures 5 and Figure 10 demonstrate example alignments for images of unhealthy retinas.

Measuring performance requires a means of validation, preferably ground truth. Manually generated ground truth is extremely difficult for such a large data set, and this is less accurate than automatic registration anyway. Fortunately, we were able to develop a multi-part alternative strategy to give pseudo ground-truth set results:

1. For the set of images from any given eye, we jointly align all images, including pairs that have little or no overlap, using the joint, multi-image algorithm [14] referenced in Section 2.3. (The current implementation of this algorithm uses Dual-Bootstrap ICP as its core pairwise registration engine.) This results in quadratic transformations between all pairs of images, even ones that failed pairwise registration. We were prepared to manually generate constraints to ensure that all images were included, but this was not necessary. No image in our data set was left out by this technique! Therefore, failures of Dual-Bootstrap ICP registration do not bias these pseudo ground-truth results.

2. We manually validated the mosaics resulting from the quadratic transformations. Several individuals, including one of the authors and a graduate student unfamiliar with the al-

gorithm, manually validated the mosaics. The image presentation enhanced the vessels so that misalignments would appear obvious.

3. As further validation we defined a notion of a "questionable registration pair". This is a pair of images (1) for which the registration CEM when restricted to one of the four image quadrants (with the central area of the image excluded) was above 2.0 pixels, or (2) a significant fraction of the vessel segments (as opposed to individual trace points) were mismatched. 75 image pairs meeting one of these conditions were called out for further analysis by a retina specialist who verified the algorithm's decision in each case.

These pseudo ground-truth results also form the next validation step: developing approximate upper bounds on the performance of point-based registration. Taking the set of vascular landmarks and centerline points for each image as given and fixed, we ask the question, "what is the best possible performance of an registration algorithm using centerline constraints?" Referring back to the objective function in Equation 1, for any pair of images we can start from the "correct" transformation and therefore find an excellent approximation to the correct set of correspondences (again, with the point sets fixed). From there we can determine the covariance of the transformation estimate. If the condition number of this matrix indicates that the transformation is sufficiently stable, we say that a point-based registration between these image pairs is possible. Denoting these pairs as $\mathcal{S}_h$ and $\mathcal{S}_p$ for the healthy and pathology sets respectively, we can measure the success rate of our algorithm as a percentage of the sizes of these two sets. This is our first performance bound. Our second, and tighter bound, restricts $\mathcal{S}_h$ and $\mathcal{S}_p$ by eliminating image pairs that have no common landmarks. We can discover this by using the "correct" transformations to find corresponding landmarks. We refer to the reduced sets as $\mathcal{S}'_h$ and $\mathcal{S}'_p$. Success rates on these sets separate the performance of initialization from the performance of the iterative minimization of the Dual-Bootstrap procedure algorithm and gives an idea of how well Dual-Bootstrap ICP does given a reasonable starting point. The cardinalities of these sets are $|\mathcal{S}_h| = 5,753$, $|\mathcal{S}_p| = 369$, $|\mathcal{S}'_h| = 5,611$, and $|\mathcal{S}'_p| = 361$.

4.2 Overall Performance

The first and most important quantitative measure of overall performance is the success rate — the percentage of image pairs for which a correct (transformation is within 1.5 pixels of the pseudo ground-truth) transformation estimate is obtained. This is summarized in Table 6 for the healthy-eye and pathology-eye datasets. These numbers are extremely high, and show virtually flawless performance of the overall registration algorithm. (By comparison, our previous retinal image registration algorithm [15] obtained 67.1% performance for the healthy-eye set, $\mathcal{S}_h$.) The few failures are due to having few common landmarks or a combination of sparse centerline trace points and low overlap. This is illustrated using a bar chart in Figure 11. To reinforce this, for image pairs that overlap in at least 35% of the pixels and have at least one correspondence, *there were no failures*. This involved over 4000 image pairs. The average CEM for accepted transformations was 0.64 pixels.

As an aside, the apparently counter-intuitive result that the pathology data set has higher success rate is explained by the pathology image pairs having higher overlap, on average. The healthy-eye images were deliberately taken to obtain complete views of the fundus, whereas the pathology-eye images were taken to capture the diseased region(s).

As a final indication of the overall performance, here is a summary of some additional experimental details:

- Using matching of single landmarks between images resulted in a 96.7% success rate, whereas matching pairs of landmarks from each image resulted in a 90.4% success rate. Since the overall performance was 97.0%, the combination of both did improve performance, although single landmark matching alone was nearly as effective.

- Over the entire dataset, including both healthy and pathology eye images, the median number of matches tried before the algorithm succeeded was 1 and the average was 5.5. The large difference between the median and the average is caused by a small number of image pairs that required an extremely large number of initial estimates before success. The worst was 746. The execution time required by the algorithm varied considerably with the number of initial estimates required before success. On a 933MHz Pentium III

computer running FreeBSD, the median time was 5 seconds.

- In a sensitivity analysis, we studied values of β (Section 3.6) in the range of 0.25 to 8. There was no change in performance for small β and only 0.4% degradation at the high end. Similarly, halving α (Section 3.3) to 5 produced no change and doubling α to 20 produced only a 0.4% degradation.

- In preliminary experiments on multimodal registration of red-free images and fluorescein angiograms, the Dual-Bootstrap ICP has nearly the same performance as the results reported. It fails for extreme cases of retinas where leakage of the dye from the blood vessels is immediate and completely obscures the vasculature, or for angiograms taken long after the injection of the dye, when the dye has slowly perfused the entire retina. In both cases, failure is largely due to a lack of initial landmarks. Figure 2 shows an example of multimodal registration.

4.3 Evaluation of the Dual-Bootstrap Procedure

Given the nearly flawless performance of our retinal image registration algorithm, the crucial issue is how much of it is due to each of the three major components of the dual-bootstrap procedure: region growth, model selection, and robust estimation. We can address this by removing each component in turn. The results are summarized in Table 3 and discussed as follows.

Region growth: Removing bootstrap region growth means that the initial similarity estimate was used to generate image-wide correspondences, as in standard ICP, and then the algorithm was run with no bootstrap region growth. The success rates were 89.4% and 82.4%. The drop is most significant — 16% — in the pathology set.

Model selection: When bootstrap model selection was eliminated, a single model was used for the entire process of bootstrap region growth and robust ICP refinement. The natural model to use is the quadratic. The first set of quadratic parameters was estimated from the correspondences in the initial bootstrap region. Using the quadratic model only led to a low success rate, as shown in Table 3. On the other hand, when we initialized an intermediate model — the reduced quadratic — from the initial bootstrap region, allowed the algorithm

to run to convergence, and then switched to the quadratic transformation, performance was much better: 94.1% on the healthy-eye set and 94.6% on the pathology-eye set. This is a heuristic form of the dual-bootstrap procedure.

Robust estimation: At first thought, the dual-bootstrapping process might seem to eliminate the need for robustness in the actual registration estimation technique. This is not true, of course. To illustrate, by simply replacing the Beaton-Tukey ρ function with a least-squares ρ function, the performance became dramatically worse (Table 3). This is because mismatches are still clearly possible. Finally, further experiments showed that the use of MUSE scale estimator over a more common estimator such as median absolute deviation [63] improved the effectiveness of the overall algorithm (97.0% vs. 93.3% on healthy eyes and 97.8% vs. 88.3% on pathology eyes).

Clearly, these experiments show that all components of Dual-Bootstrap ICP are important, with importance increasing substantially for the more difficult pathology eye data set.

4.4   Comparison to Other Retinal Image Registration Algorithms

Section 2.2 summarized the literature on retinal image registration and discussed some of the limitations of current techniques. Some published techniques do report numerical results. Many of these only report fewer than 10 results, so these are not considered carefully here. More details are given in [4, 15, 62]. In particular, [4] reports accuracies of 1.35 pixels for 180 video resolution images, while [62] reports sub-pixel accuracy in registering images centered on the optic disk, but only with a large trade-off in the number of function evaluations. Only in recent work [15] are more that 100 image pairs evaluated and performance quantified as a function of the overlap between images. As seen in Figure 11 and discussed in Section 4.2, the Dual-Bootstrap ICP substantially outperforms [15].

5   Discussion and Conclusions

We have introduced the Dual-Bootstrap ICP algorithm and successfully applied it to retinal image registration. The idea behind the algorithm is to start from an initial estimate that is

only assumed to be accurate over a small region. Using the covariance matrix of the estimated transformation parameters as its guide, the approach bootstraps both the region over which the model is applied and the choice of transformation models. A robust version of standard ICP is applied to re-estimate the actual transformation in each bootstrap region. The implementation uses an automatic initialization technique based on matching single vascular landmarks from each image or matching pairs of vascular landmarks from each image.

Experimental evaluation showed nearly flawless performance of the Dual-Bootstrap ICP algorithm on a large set of retinal image pairs. In comparison to current retina image registration algorithms, it handles lower image overlaps, image changes and poor image quality, all of which reduce the number of common landmarks between images. Moreover, by effectively exploiting the vascular structure during the dual-bootstrap procedure it avoids the need for expensive global search techniques. All of this is accomplished automatically without the need for any significant tuning parameters.

Although the experimental results reported only apply Dual-Bootstrap ICP to retinal image registration, the algorithm design, especially the dual-bootstrap procedure, is intended for broad applicability. We have already begun applying the algorithm to 3D registration of vascular structures and neural structures using confocal microscopy. The extension to these is straightforward, and the results are excellent thus far. Extension to other domains is possible as well. The keys to doing this are developing a method of initialization based on small image regions and developing a series of transformation models.

The remainder of this discussion concentrates on intuitions about the advantages and applicability of the Dual-Bootstrap ICP algorithm in general:

- In comparison with current indexing-based initialization methods and minimal-subset random-sampling methods, Dual-Bootstrap ICP has the major advantage requiring fewer initial correspondences. This is because it starts from an initial low-order transformation that must only be accurate in small initial regions.

- In comparison with multiresolution methods, Dual-Bootstrap ICP starts from what might be thought of as a key-hole view on the alignment rather than the bird's-eye view of multiresolution. An experimental algorithm using robust ICP in combination with mul-

tiresolution failed to align some of the more difficult retinal image pairs.

- Instead of matching globally, which could require simultaneous consideration of multiple matches [19], Dual-Bootstrap ICP uses region and model bootstrapping to resolve matching ambiguities. We briefly experimented with the use of multiple matches per feature, and found that the dual-bootstrap procedure made this unnecessary.

- It may be possible to use other techniques besides robust ICP to estimate the transformation parameters within the bootstrap region. This is especially important when there is no clear relationship between features in the two images. Any effective technique can be used as long as a parameter estimate covariance matrix can be computed. At the very worst, the computation can be done using numerical differentiation.

- Deciding when the dual-bootstrap approach is needed requires analyzing the combination of modeling error (see Table 1) and transfer error of initial estimates. If this error, image-wide, is at least half the distance between the structures (e.g. blood vessels) alignment is based on, then the dual-bootstrap approach is needed. Interestingly, this question can be reversed to ask if, by using the dual-bootstrap approach, the initial conditions required for registration can be weakened and therefore more difficult problems addressed.

- Other than the obvious case of no valid initial correspondences, the dual-bootstrap approach can fail in two ways. The first is when the initial model is too weak. One example might be use of an initial rigid transformation when significant scaling is present. The second way it can fail is when the images contain two geometrically-separated clusters of features, and the initial transformation is estimated only in a single cluster. As the gap between clusters grows, the transfer error will grow with it, potentially leading to mismatches when the bootstrap region R_t grows to include a second cluster.

As a final remark, in the context of retinal image registration, the success of the Dual-Bootstrap ICP algorithm, somewhat counter-intuitively, has thrown the research challenge back to feature extraction. The algorithm so successfully exploits whatever data are available that truly the only cause of failure is extremely poor quality image data, leading to an extremely large

fraction of missed or spurious features. Thus, developing robust, low-contrast feature extraction is our main avenue to further improvements in retinal image registration.

6 Acknowledgments

References

[1] K. Al-Kofahi, S. Lasek, S. D., C. Pace, G. Nagy, J. N. Turner, and B. Roysam. Rapid automated three-dimensional tracing of neurons from confocal image stacks. *IEEE Transactions on Information Technology in Biomedicine*, 6(2):171–187, 2002.

[2] S. Aylward and E. Bullitt. Initialization, noise, singularities, and scale in height-ridge traversal for tubular object centerline extraction. *IEEE Transactions on Medical Imaging*, 21:61–75, 2002.

[3] S. Aylward, J. Jomier, S. Weeks, and E. Bullitt. Registration of vascular images. *International Journal of Computer Vision*, (to appear) 2003.

[4] 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.

[5] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Proceedings of the Second European Conference on Computer Vision*, pages 237–252, 1992.

[6] J. Berkow, R. Flower, D. Orth, and J. Kelley. *Fluorescein and Indocyanine Green Angiography, Technique and Interpretation*. American Academy of Ophthalmology, 2nd edition, 1997.

[7] 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.

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

[9] G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, 34(3):344–371, June 1986.

[10] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, December 1992.

[11] K. Bubna and C. V. Stewart. Model selection techniques and merging rules for range data segmentation algorithms. *Computer Vision and Image Understanding*, 80:215–245, 2000.

[12] 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 Transactions on Information Technology in Biomedicine*, 3(2):125–138, 1999.

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

[14] 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.

[15] 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.

[16] G. Champleboux, S. Lavallee, R. Szeliski, and L. Brunie. From accurate range imaging sensor calibration to accurate model-based 3-d object localization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 83–89, 1992.

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

[18] C. Chua and R. Jarvis. 3d free-form surface registration and object recognition. *International Journal of Computer Vision*, 17(1):77–99, 1996.

[19] H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages II:44–51, 2000.

[20] A. V. Cideciyan. Registration of ocular fundus images. *IEEE Engineering in Medicine and Biology*, 14(1):52–58, Jan. 1995.

[21] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. Registration of 3d multi-modality medical images using surfaces and point landmarks. *Pattern Recognition Letters*, 15(5):461–467, 1994.

[22] P.-E. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.

[23] M. DeGrezia and M. Robinson. Ophthalmic manifestations of HIV: an update. *Journal of the Association of Nurses in AIDS Care*, 12(3):22–32, May-Jun 2001.

[24] Diabetic Retinopathy Research Group. Diabetic retinopathy study. Report Number 7. A modification of the Airlie House classification of diabetic retinopathy. *Investigative Ophthalmology & Visual Science*, 21:210–??, Dec 1981.

[25] J. S. Duncan and N. Ayache. Medical image analysis: progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):85–105, 2000.

[26] B. Ege, T. Dahl, T. Sndergaard, O. Larsen, T. Bek, and O. Hejlesen. Automatic registration of ocular fundus images. In *Workshop on Computer Assisted Fundus Image Analysis*, May 2000.

[27] J. Feldmar, J. Declerck, G. Malandain, and N. Ayache. Extension of the ICP algorithm to nonrigid intensity-based registration of 3d volumes. *Computer Vision and Image Understanding*, 66(2):193–206, May 1997.

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

[29] K. Fritzsche, A. Can, H. Shen, C. Tsai, J. Turner, H. Tanenbuam, C. Stewart, and B. Roysam. Automated model based segmentation, tracing and analysis of retinal vasculature from digital fundus images. In J. S. Suri and S. Laxminarayan, editors, *State-of-The-Art Angiography, Applications and Plaque Imaging Using MR, CT, Ultrasound and X-rays*. Academic Press, 2002.

[30] W. Grimson, T. Lozano-Perez, W. Wells, G. Ettinger, and S. White. An automatic registration method for frameless stereotaxy, image, guided surgery and enhanced reality visualization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 430–436, 1994.

[31] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, editors. *Medical Image Registration*. CRC Press, 2001.

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

[33] W. Hart and M. Goldbaum. Registering retinal images using automatically selected control point pairs. In *Proceedings IEEE International Conference on Image Processing*, volume 3, pages 576–581, 1994.

[34] R. Hartley and A. Zisserman. *Multiple View Geometry.* Cambridge University Press, 2000.

[35] C. Heneghan, P. Maguire, N. Ryan, and P. de Chazal. Retinal image registration using control points. *IEEE International Symposium on Biomedical Imaging*, pages 349 –352, July 2002.

[36] K. Higuchi, M. Hebert, and K. Ikeuchi. Building 3-d models from unregistered range images. *Graphical Models and Image Processing*, 57(4):315–333, 1995.

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

[38] A. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000.

[39] R. Jagoe, C. Blauth, P. Smith, J. Arnold, K. Taylor, and R. Wootton. Automatic geometrical registration of fluorescein retinal angiograms. *Computers and Biomedical Research*, 23(5):403–9, Oct 1990.

[40] A. Johnson and M. Hebert. Surface matching for object recognition in complex 3-dimensional scenes. *Image and Vision Computing*, 16(9-10):635–651, July 1998.

[41] H. Lester and S. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32(1):129–149, 1999.

[42] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The Digital Michelango Project: 3d scanning of large statues. In *Computer Graphics Proceedings (SIGGRAPH)*, pages 131–144, New Orleans, July 2000.

[43] D. Lloret, J. Serrat, A. Lopez, A. Soler, and J. Villanueva. Retinal image registration using creases as anatomical landmarks. In *Proceedings IEEE International Conference on Image Processing*, volume 3, pages 203–207, 2000.

[44] D. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31(3):355–395, 1987.

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

[46] 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.

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

[48] G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita. Automatic retinal image registration scheme using global optimization techniques. *IEEE Transactions on Information Technology in Biomedicine*, 3(1):47–60, 1999.

[49] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[50] A. Mendonca and J. Campilho, A. Nunes. A new similarity criterion for retinal image registration. In *Proceedings IEEE International Conference on Image Processing*, pages 696–700, 1994.

[51] C.-H. Menq, H.-T. Yau, and G.-Y. Lai. Automated precision measurement of surface profile in CAD-directed inspection. *IEEE Transactions on Robotics and Automation*, 8(2):268–278, 1992.

[52] J. V. Miller. *Regression-Base Surface Reconstruction: Coping with Noise, Outliers, and Discontinuities.* PhD thesis, Rensselaer Polytechnic Institute, Aug. 1997.

[53] J. V. Miller and C. V. Stewart. MUSE: Robust surface fitting using unbiased scale estimates. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 300–306, 18–20 June 1996.

[54] N. Mourvliansky, G. Matsopoulos, K. Delibasis, and K. Nikita. Automatic retinal registration using global optimization techniques. In *Proceedings of the 20th IEEE International Conference Engineering in Medicine and Biology*, volume 2, 1998.

[55] J. Mundy and A. Zisserman, editors. *Geometric Invariance in Computer Vision*. MIT Press, 1992.

[56] V.-D. Nguyen, V. Nzomigni, and C. Stewart. Fast and robust registration of 3d surfaces using low curvature patches. In *Proceedings of the 2nd International Conference on 3-D Digital Imaging and Modeling*, pages 201–208, 1999.

[57] J. Park, J. Keller, P. Gader, and R. Schuchard. Hough-based registration of retinal images. In *1998 IEEE International Conference on Systems, Man, and Cybernetics*, volume 5, pages 4550–4555, 1998.

[58] E. Peli, R. A. Augliere, and G. T. Timberlake. Feature-based registration of retinal images. *IEEE Transactions on Medical Imaging*, 6(3), September 1987.

[59] G. P. Penny, J. Weese, J. A.Little, P. Desmedt, D. L. Hill, and D. J. Hawkes. A comparison of similarity measures for used in 2-d-3-d medical image registration. *IEEE Transactions on Medical Imaging*, 17(4):586–594, 1998.

[60] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger. Mapping the human retina. *IEEE Transactions on Medical Imaging*, 17(4):606–620, Aug 1998.

[61] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1992.

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

[63] P. J. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79:871–880, 1984.

[64] P. J. Rousseeuw and C. Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88:1273–1283, 1993.

[65] P. J. Saine and M. E. Tyler. *Ophthalmic Photography*. Butterworth Heinemann, 2002.

[66] H. Sawhney, S. Hsu, and R. Kumar. Robust video mosaicing through topology inference and local to global alignment. In *Proceedings of the Fifth European Conference on Computer Vision*, volume II, pages 103–119, 1998.

[67] H. Sawhney and R. Kumar. True multi-image alignment and its application to mosaicing and lens distortion correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(3):235–243, 1999.

[68] C. Schmid, R. Mohr, and C. Bauckhage. Comparing and evaluating interest points. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 230–235, 1998.

[69] H. Shen, C. Stewart, B. Roysam, G. Lin, and H. Tanenbaum. Frame-rate spatial referencing based on invariant indexing and alignment with application to laser retinal surgery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 79–86, 2001.

[70] H. Shen, C. Stewart, B. Roysam, G. Lin, and H. Tanenbaum. Frame-rate spatial referencing based on invariant indexing and alignment with application to laser retinal surgery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3), March 2003.

[71] H. Shum and R. Szeliski. Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, 36(2):101–130, 2000.

[72] C. Sinthanayothin, J. Boyce, H. Cook, and T. Williamson. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *The British Journal of Ophthalmology*, 83(3):902–10, Aug 1999.

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

[74] W. Tan, Y. Wang, and S. Lee. Retinal blood vessel detection using frequency analysis and local-mean-interpolation filters. In *SPIE Conference Medical Imaging:Image Processing*, volume 4322, pages 1373–1384, 2001.

[75] J. Thirion. New feature points based on geometric invariants for 3d image registration. *International Journal of Computer Vision*, 18(2):121–137, 1996.

[76] P. Torr. An assessment of information criteria for motion model selection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 47–52, 1997.

[77] P. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1):271–300, 2002.

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

[79] P. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, April 2000.

[80] C.-L. Tsai, C. Stewart, B. Roysam, and H. Tanenbaum. Covariance-driven retinal image registration initialized from small sets of landmark correspondences. *IEEE International Symposium on Biomedical Imaging*, July 2002.

[81] 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*, to appear 2003.

[82] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

[83] R. Weleber and K. Gregory-Evans. *Retina*, chapter 18. Retinitis Pigmentosa and Allied Disorders, pages 362–460. Mosby, 2001.

[84] J. Yu and C. Hung B.N. Liou. Fast algorithm for digital retinal image alignment. In *IEEE Engineering in Medicine and Biology Society, 11th Annual International Conference*, pages 0734–0735, 1989.

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

[86] F. Zana and J.-C. Klein. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, 10(7):1010–1019, 2001.

[87] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.

[88] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2):161–195, 1998.
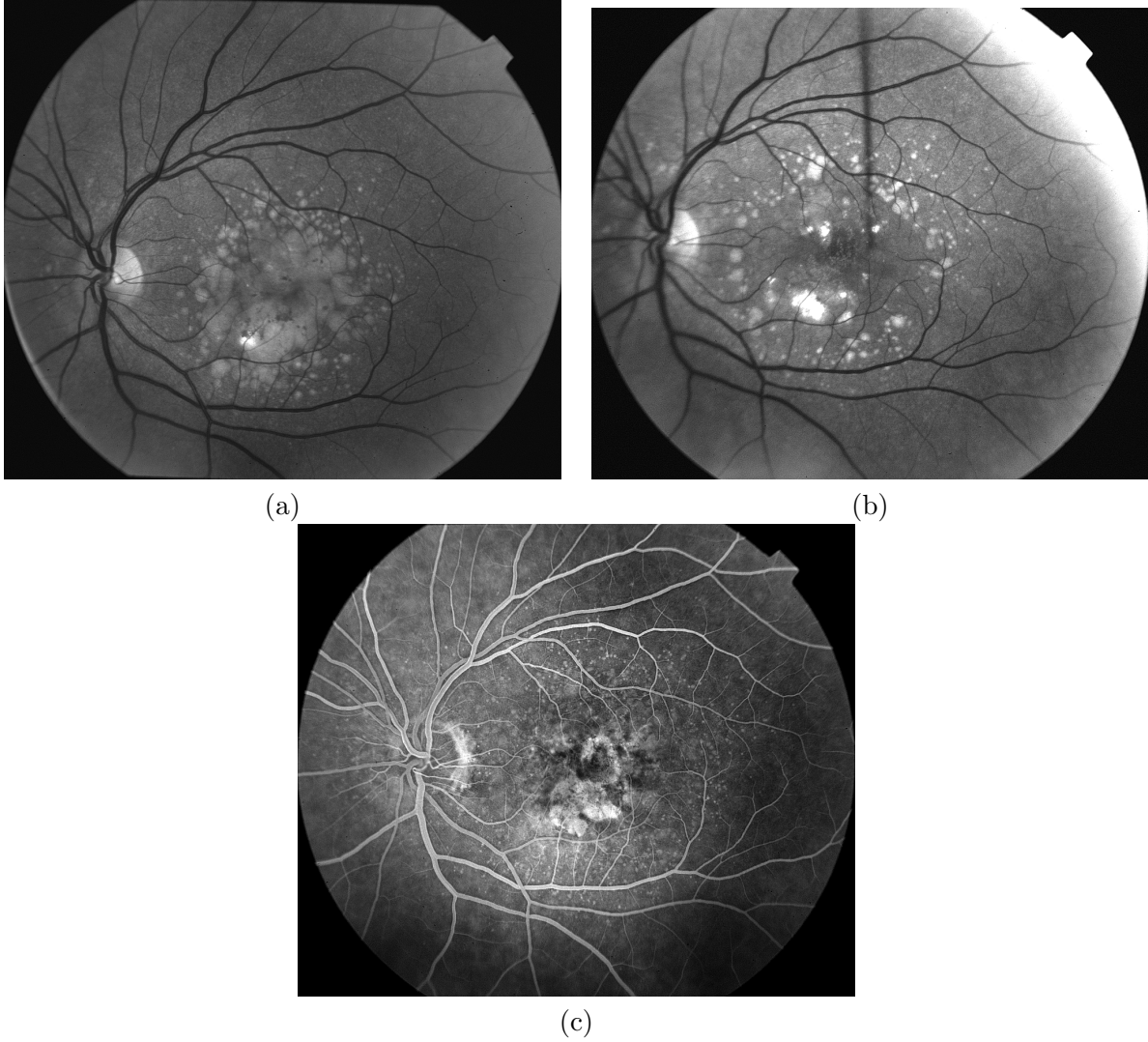
Figure 1: Images (a) & (b) are red-free images of a patient's retina taken 3 years apart. Differences between the images are caused by the progression of non-exudative Age-related Macular Degeneration (AMD). Image (c) is a fluorescein angiogram image of the same patient taken on the same date as (b).
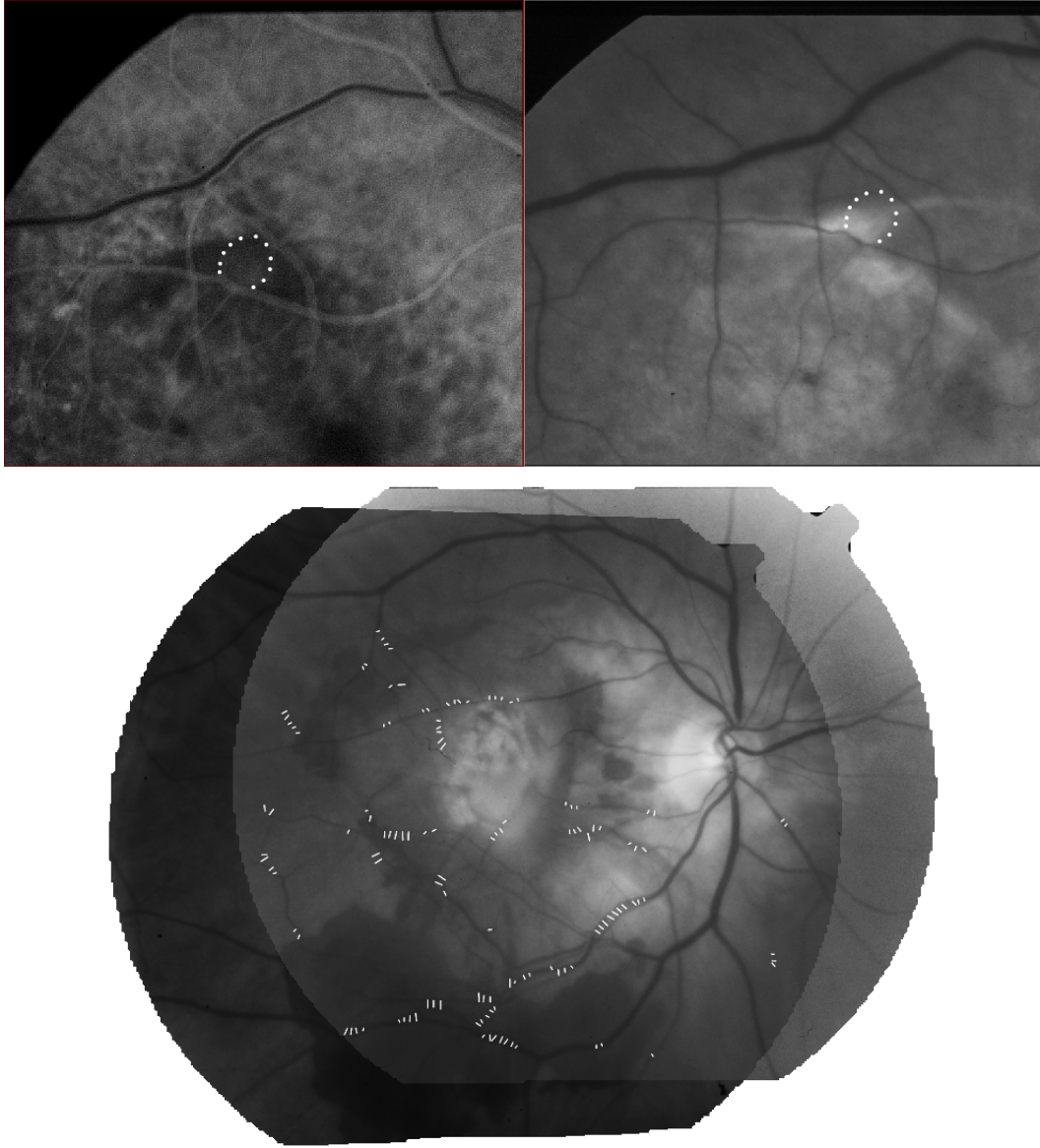
Figure 2: Applications of retinal image registration. The top image pair shows cropped regions of an aligned fluorescein angiogram (left) and red-free image (right). On the angiogram, an ophthalmologist has highlighted leakage of the fluorescein dye. The same region is automatically highlighted on the red-free image, which shows what the ophthalmologist sees when looking through the lens of the eye at the retina. The bottom panel shows the alignment of two images of a patient having exudative AMD. The images were taken 2.5 years apart. Shifts in the position of the vasculature of more than 4 pixels are highlighted with line segments.
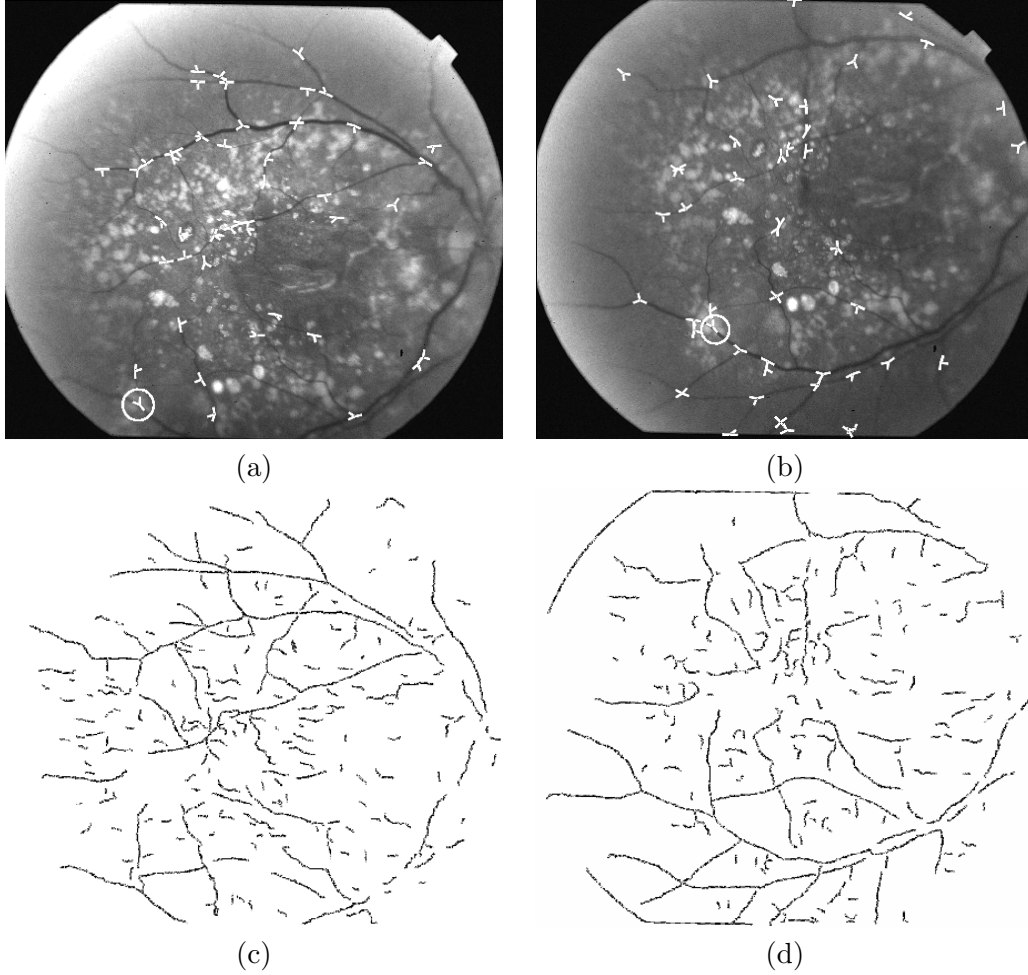
Figure 3: Fundus images of an unhealthy eye (nonexudative AMD) together with image features used in Dual-Bootstrap ICP registration. Panels (a) and (b) show the images, with landmarks extracted by our retinal image tracing algorithm [12, 29]. The landmarks are branching and crossover points of the retinal vasculature. The only landmark that is common to the two images is circled. Panels (c) and (d) show the centerline points obtained by the tracing algorithm. Many inconsistencies in the two sets of traces may be observed.

(a)   (b)
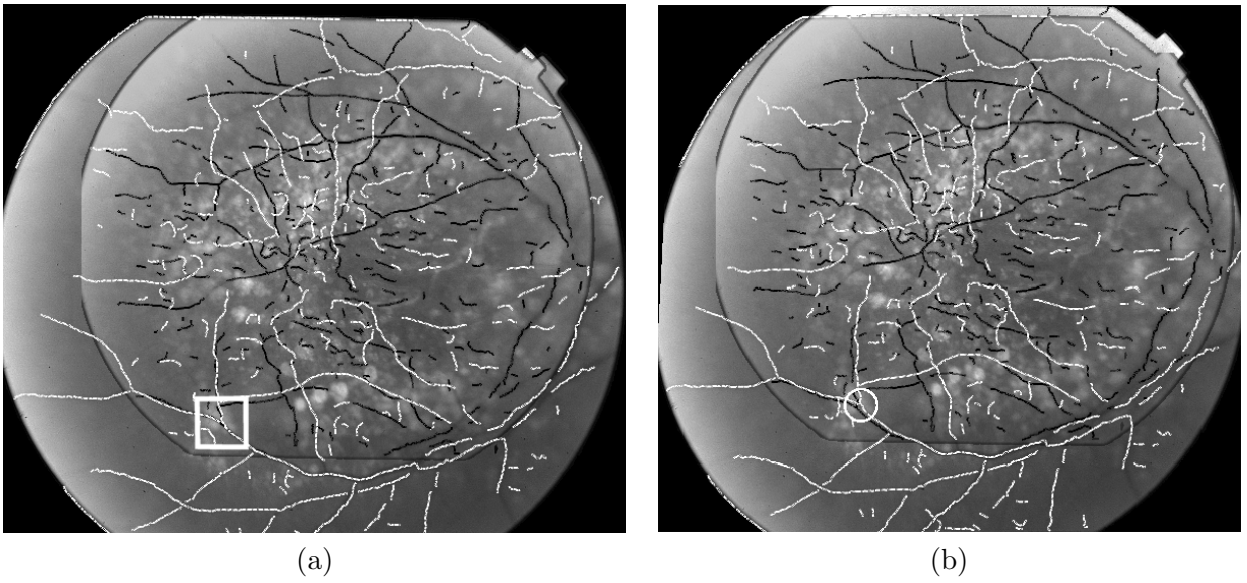
Figure 4: The results of iteratively aligning the images shown in Figure 3. Panel (a) shows the initial alignment based on the single common landmark. Trace points from Figure 3c are shown in black and trace points from Figure 3d are shown in white. The white square on the bottom left shows the accurate alignment surrounding the initial correspondence. Panel (b) shows the alignment result after convergence of the robust ICP alignment process. The alignment is extremely poor due to the mismatches in the upper region.
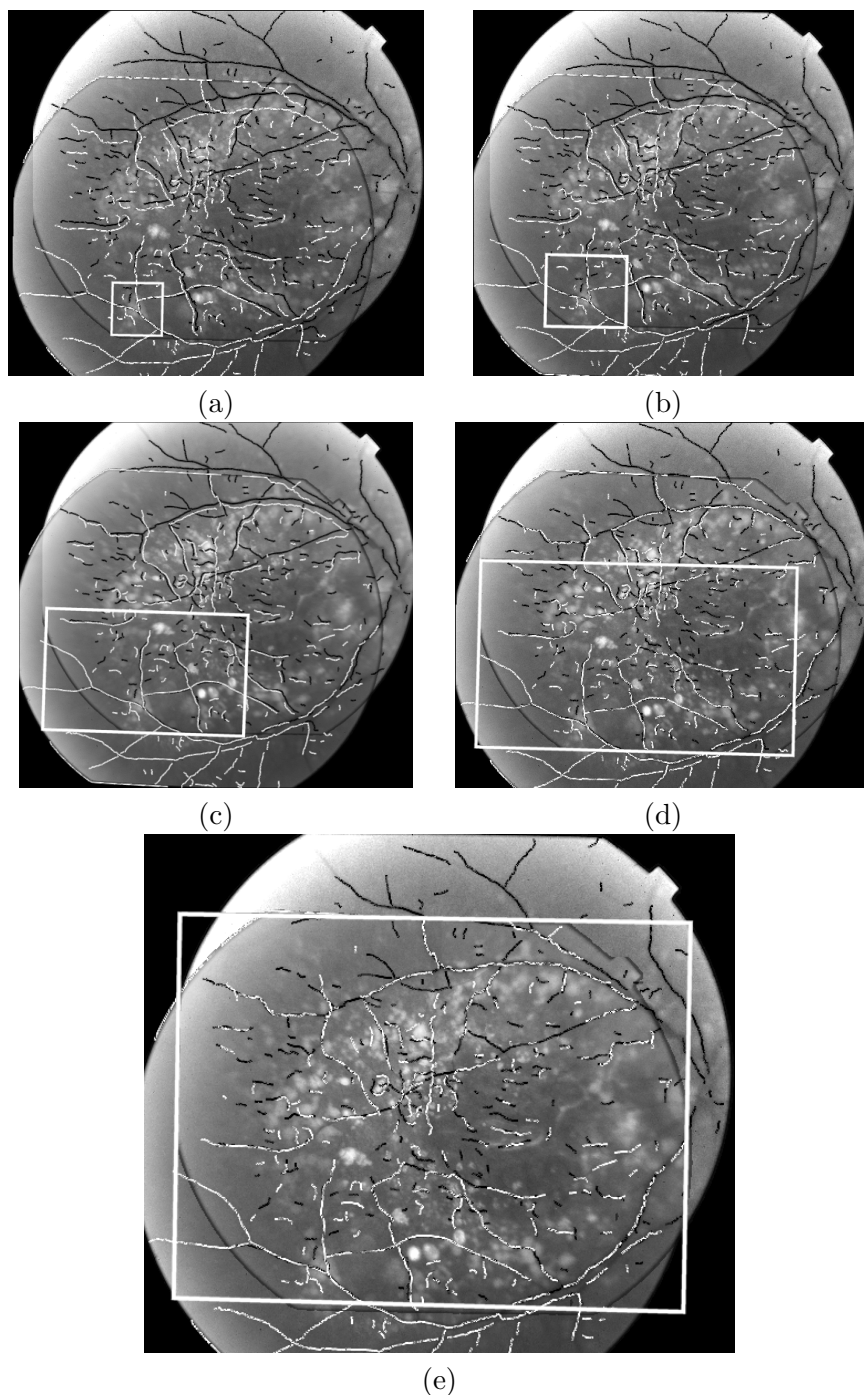
Figure 5: Illustrating the Dual Bootstrap ICP algorithm in retinal image registration using the images and feature points shown in Figure 3 and the initial alignment shown in Figure 4a. In each iteration, a robust version of ICP is applied *only* in the bootstrap region, R_t, indicated by the white rectangle in each figure. The transformation is only required to be accurate in this bootstrap region. Also in each iteration, the best transformation model (in this case, similarity, reduced-quadratic, or quadratic — see Table 1) — is automatically selected *and* the bootstrap region is grown. Several increments of this process are shown in the panels. In panels (a)-(c) Dual-Bootstrap ICP selected a similarity transformation. In panel (d) it selected a reduced-quadratic. Panel (e) shows the final alignment using a 12-parameter quadratic model.

Transformation	Equation	DoF	Accuracy
Similarity	$\mathbf{p}' = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & 0 & 0 & 0 \\ \theta_{21} & -\theta_{13} & \theta_{12} & 0 & 0 & 0 \end{pmatrix} \mathbf{X}(\mathbf{p} - \mathbf{p_0})$	4	5.05 pixels
Affine	$\mathbf{p}' = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & 0 & 0 & 0 \\ \theta_{21} & \theta_{22} & \theta_{23} & 0 & 0 & 0 \end{pmatrix} \mathbf{X}(\mathbf{p} - \mathbf{p_0})$	6	4.58 pixels
Reduced quadratic	$\mathbf{p}' = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & 0 & \theta_{14} \\ \theta_{21} & -\theta_{13} & \theta_{12} & \theta_{24} & 0 & \theta_{24} \end{pmatrix} \mathbf{X}(\mathbf{p} - \mathbf{p_0})$	6	2.41 pixels
Quadratic	$\mathbf{p}' = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{pmatrix} \mathbf{X}(\mathbf{p} - \mathbf{p_0})$	12	0.64 pixels

Table 1: The set of transformation models used in retinal image registration. To clarify notation in the equations, $\mathbf{p} = (x, y)^T$ is an image location in I_1, $\mathbf{p_0} = (x_0, y_0)^T$ is the center of the bootstrap region in image I_1, $\mathbf{X}(\mathbf{p} - \mathbf{p_0}) = (1, x - x_0, y - y_0, (x - x_0)^2, (x - x_0)(y - y_0), (y - y_0)^2)^T$ is the vector of monomial terms of the quadratic basis, and $\mathbf{p}'$ is the transformed location in the coordinate system of I_2. The centering is required in the reduced quadratic model, but is used in practice in all models for numerical reasons. In addition to the formulations, the table also shows the degrees of freedom (DoF) in each model and the average alignment error on 1024×1024 images.
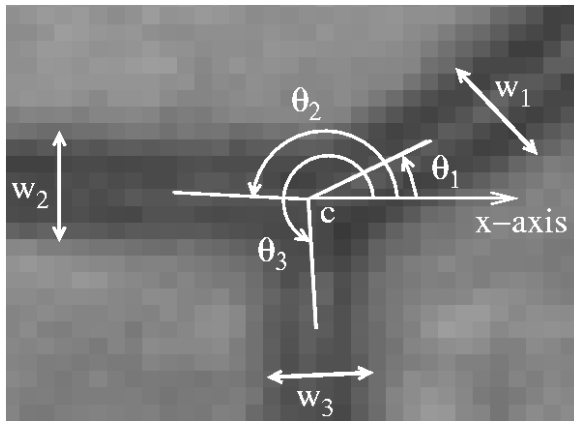
Figure 6: A landmark is characterized by a center location **c**, the orientations of the three blood vessels that meet to form it and the widths w_j of the blood vessel. Differences in orientations and the ratios of the blood vessel widths are invariant to rotation, translation and scaling of the image. The orientations themselves are invariant to translation and scaling.
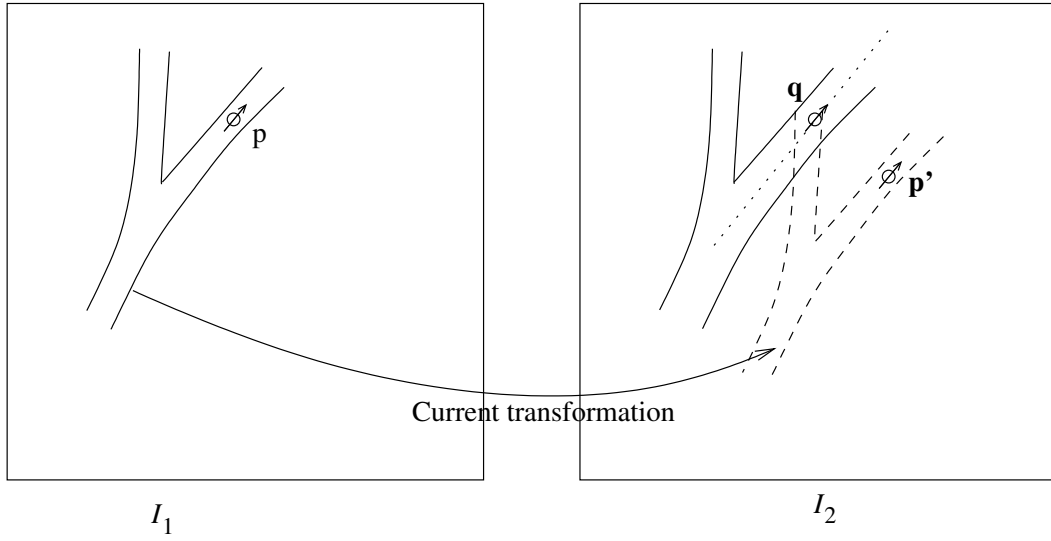
Figure 7: A schematic illustrating trace point matching and the distance metric. Trace point $\mathbf{p}$ from I_1 is mapped into I_2 based on the current transformation estimate (point $\mathbf{p}'$), and the closest trace $\mathbf{q}$ in I_2 is found. (The small arrows represent the trace directions.) The error distance is the distance from $\mathbf{p}'$ to the line tangent to the trace through $\mathbf{q}$ (the dotted line). Letting $\boldsymbol{\eta}_q$ be the normal to the line, the distance is $|(\mathbf{p}' - \mathbf{q})^T \boldsymbol{\eta}_q|$.
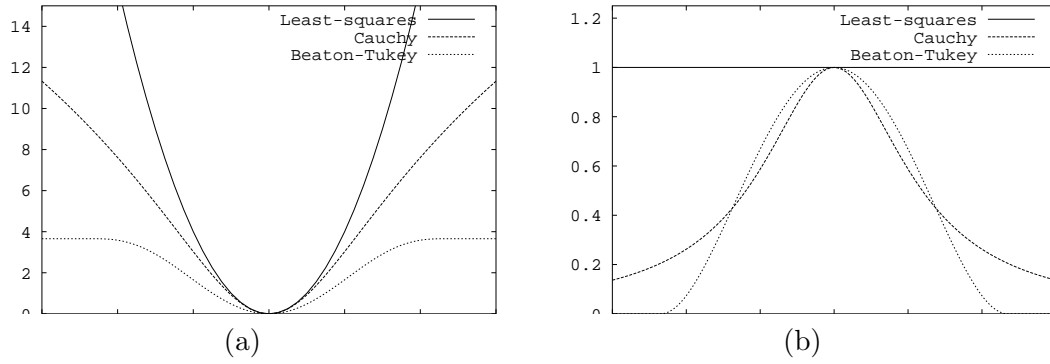
Figure 8: Plots of (a) the robust loss function $\rho(u)$ and (b) weight function $w(u) = \rho'(u)/u$ for the Beaton-Tukey biweight loss function, the Cauchy loss function and the quadratic loss function, which equates to least-squares estimation. The Beaton-Tukey is chosen because it most aggressively rejects outliers, providing no weight to matches with normalized distances greater than about 4 standard deviations.
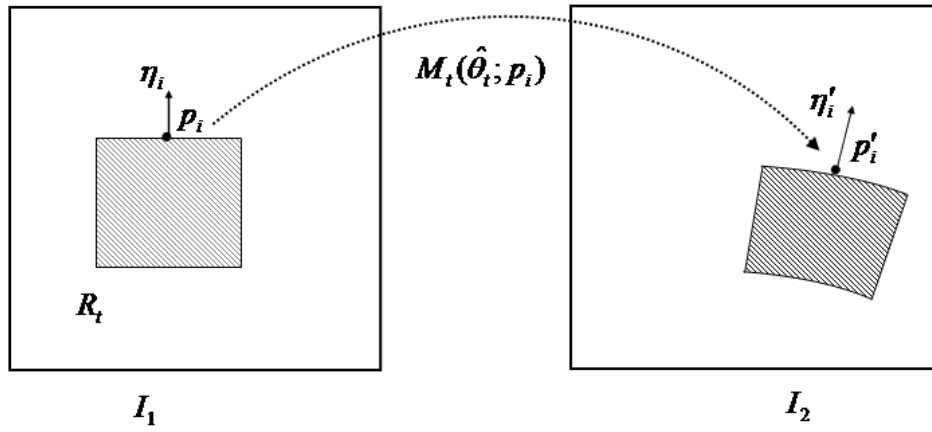
Figure 9: The bootstrap region, R_t, is expanded by moving out perpendicular to each side in image I_1. The region and its mapping (never explicitly computed) onto I_2 are shaded. Point $\mathbf{p}_i$ is in the center of one side of R_t and $\boldsymbol{\eta}_i$ is the outward normal to R_t at $\mathbf{p}_i$. The transfer error of mapping $\mathbf{p}_i$ onto I_2 helps determine the growth of the bootstrap region.
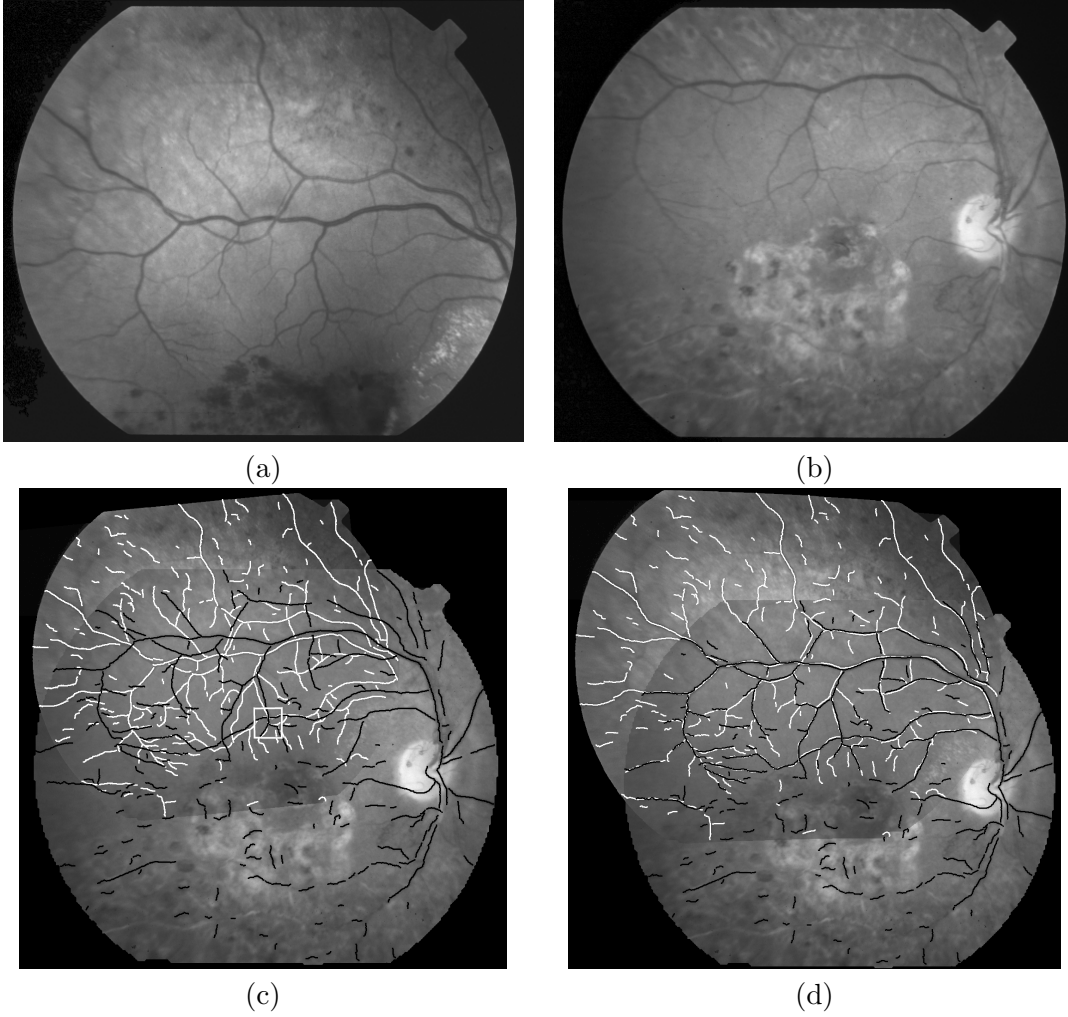
Figure 10: Dual Bootstrap ICP retinal image registration results on a pair of images taken 3.5 years apart. When the image shown in (b) was taken, the patient had a branch vein occlusion. Substantial changes in the non-vascular regions appear. Panel (c) shows the alignment of the two images based on the initial bootstrap region. Extracted blood vessel centerline points from the two different images are shown in black and in white. The initial bootstrap region is shown as a white box. Clearly the alignment is quite poor image-wide. Panel (d) shows the final alignment estimated by the Dual-Bootstrap ICP algorithm starting from this initial estimate.

	all pairs	one landmark pairs
healthy — $\mathcal{S}_h$ (%)	97.0	99.5
pathology — $\mathcal{S}_p$ (%)	97.8	100

Table 2: Overall success rate of the Dual-Bootstrap ICP retinal image registration algorithm on healthy-eye and pathological-eye images. The first column, labeled "all pairs", is for all "correct" image pairs — the sets $\mathcal{S}_h$ and $\mathcal{S}_p$. The second column, labeled "one landmark", is for all "correct" image pairs having at least one common landmark — the sets $\mathcal{S}'_h$ and $\mathcal{S}'_p$.
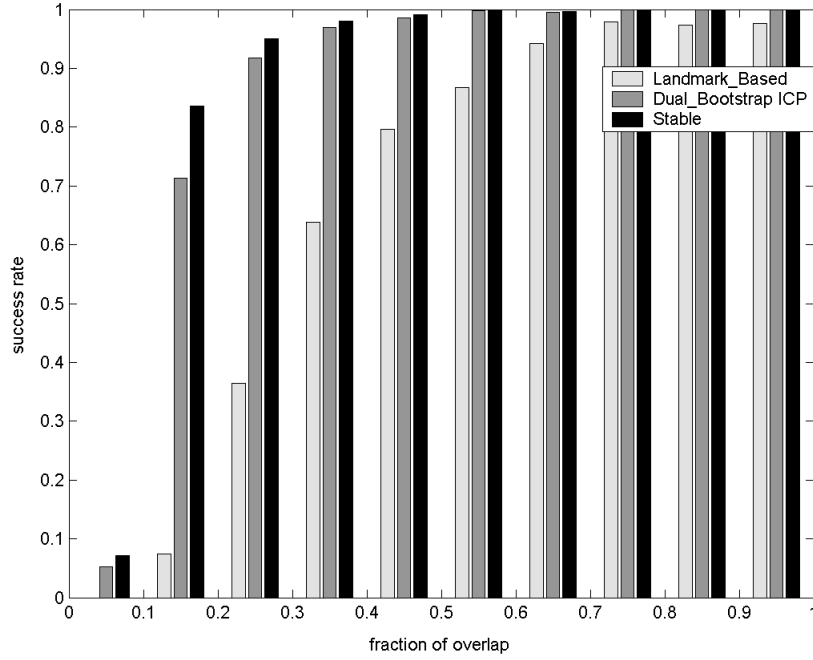
Figure 11: Plotting the percentage of successful retinal image registrations as a function of overlap between images. The plots include all image pairs in healthy-eye data set $\mathcal{S}_h$, not just those for which a stable transformation is available. The percentage for the Dual-Bootstrap ICP based algorithm, the percentage for our earlier landmark-based algorithm [15], and the percentage of stable transformations are all shown for each interval. When the height for an algorithm equals that of the height of "stable", 100% success was obtained by the algorithm. Plotting the results this way shows the overall difficulty in obtaining enough information to register at extremely low overlaps. Even here, however, the success rate of the Dual-Bootstrap ICP based algorithm nearly matches the best possible for a fixed set of trace points.

	No region growth	No model selection	No robustness
Healthy $\mathcal{S}_h$ (%)	89.4	84.7	39.0
Pathology $\mathcal{S}_p$ (%)	82.4	80.5	12.5

Table 3: Success rates of retinal image registration when each of the three main component of the Dual-Bootstrapping ICP algorithm was removed separately: region growth, model selection, and robust estimation. These are significantly lower than the 97.0% and 97.8% numbers for the overall algorithm.