accepted for publicaction in *IEEE Trans. on Biomedical Engineering*

Predictive Scheduling Algorithms for Real-time Feature Extraction and Spatial Referencing: Application to Retinal Image Sequences

Gang Lin[1] Charles V. Stewart[1] Badrinath Roysam[1] Kenneth Fritzsche[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

October 30, 2002

Abstract

Real-time spatial referencing is an important alternative to tracking for designing spatially-aware ophthalmic instrumentation for procedures such as laser photocoagulation and perimetry. It requires independent, fast registration of each image frame from a digital video stream (1024×1024 pixels) to a spatial map of the retina. Recently, we have introduced a spatial referencing algorithm that works in three primary steps: (1) tracing the retinal vasculature to extract image feature (landmarks), (2) invariant indexing to generate hypothesized landmark correspondences and initial transformations, and (3) alignment and verification steps to robustly estimate a 12-parameter quadratic spatial transformation between the image frame and the map. The goal of this paper is to introduce techniques to minimize the amount of computation for successful spatial referencing. The fundamental driving idea is to make feature extraction subservient to registration and therefore only produce the information needed for verified, accurate transformations. To this end, the image is analyzed along one-dimensional, vertical and horizontal grid lines to produce a regular sampling of the vasculature, needed for step (3) and to initiate step (1). Tracing of the vascular is then prioritized hierarchically to quickly extract landmarks and groups (constellations) of landmarks for indexing. Finally, the tracing and spatial referencing computations are integrated so that landmark constellations found by tracing are tested immediately. The resulting implementation is an order-of-magnitude faster with the same success rate. The average total computation time is 45.5 milliseconds per image on a 2.2 GHz Pentium Xeon processor.

Index terms: biomedical image analysis, retinal image sequences, feature-based image registration, real-time computing, scheduling, vasculature tracing, spatial mapping, spatial referencing, mosaic synthesis.

1 Introduction

The present work is inspired by the need to build real-time, automatic image analysis tools to assist ophthalmic procedures such as laser retinal surgery [20, 21, 23, 16], and perimetry [13]. In these procedures, the core image analysis problem is to accurately locate specific points on the retina, such as the location of a laser spot, and transform these locations into a pre-established spatial frame of reference. We term this problem "Spatial Referencing". Figure 1 illustrates spatial referencing as a feature-based image registration problem where estimation of a spatial transformation from on-line images to a pre-established spatial map is the central issue. Spatial referencing must be performed quickly enough to enable measurements such as dosimetry calculations and to enable real-time control decisions such as safety shutoffs. Frame rates of 30 milliseconds (ms) or higher are desirable.

Spatial referencing should not be approached entirely as a frame-to-frame tracking problem, despite earlier work in this area [1, 2, 24, 18]. Frequent, large-scale changes between images caused by saccadic eye movements [12, 27], blinking, and glare mean that such incremental methods can only work for relatively short durations. What is needed is the ability to locate the surgical tool (in this case, a laser beam) absolutely, without using estimates from previous frames. This requires a reliable spatial reference. Such a capability could be combined with short-duration frame-to-frame tracking, as in the earlier work of Becker, [2]. Our focus here, however, is on the absolute spatial reference.

In a recent work [29, 31, 32] we have described a new approach to spatial referencing in which each image frame acquired on-line is registered to a spatial map of the retina pre-computed from diagnostic images (Figure 1). The on-line image dimensions are 1024×1024 pixels and the geometric transformation model mapping images onto the spatial map has 12 parameters [8]. Together, these present special challenges when trying to meet frame-rate computational deadlines. Our approach is based on a combination of techniques. Features used in registration are extracted from the distinctive and stable vascular pattern of the retina using a fast, exploratory algorithm [6, 30, 35]. Quasi-invariant indexing is used to generate hypothesized correspondences between on-line image features and features computed from off-line diagnostic images. A hierarchical, robust verification and refinement algorithm is used to convert initial transformation estimates that use low-order models into the image-wide, 12-parameter, quadratic mapping [8]. This method is described in more detail in Section 2 below.

In building on this algorithm to develop a reliable real-time spatial referencing system, several issues must be addressed. The two most important are (1) making the core algorithms as efficient and effective as possible, and (2) making the time required by the computation predictable [26, 17] so that real-time measurement and control decisions can be made. The contribution of the present work is an adaptive, opportunistic method to minimize and prioritize the feature extraction and registration computations for successful spatial referencing. A separate paper, in preparation, describes some of the real-time operating system design issues in addressing the second.

In addressing the efficiency issue, our design principle is to view the estimation and verification of the 12-parameter transformation as the only required computation, and to develop opportunistic methods to extract the minimum amount of information from the on-line image to accomplish this goal. This leads naturally to algorithms that integrate partial feature extraction methods [30], and spatial referencing computations [32]. This approach is shown to result in an order of magnitude improvement speed with no significant loss of accuracy.

A graphic illustration of a good case is provided in Figure 10. The yellow outline in the figure highlights the on-line image frame that was registered to a complete retinal mosaic. The green lines indicate the vessels in the on-line image frame that were traced. In this case, tracing was applied in just two of the small grid boxes (amounting to less than 1% of the image), but this provided enough information from tracing to find an accurate image-wide registration. The rest of this paper describes adaptive algorithms for achieving this level of efficiency.

2 Background on Retinal Image Tracing and Spatial Referencing

This section provides a summary of the feature extraction and spatial referencing algorithms. These have been reported in previous papers and form the context for the computational efficiency optimization techniques presented in this paper.

2.1 Exploratory Algorithms for Tracing Retinal Vasculature

Recently [6, 30], we have described fast exploratory algorithms for tracing the retinal vasculature, improving upon prior work [9, 10, 11, 25]. Our algorithms are based on modeling the vessels as dark elongated structures with roughly anti-parallel edges. The algorithms proceed in three stages:

1. **Grid analysis and seed point detection** surveys the image along a grid of evenly spaced pixel-wide horizontal and vertical lines through the image (Figure 2a). This estimates the local image statistics and noise levels. Blood vessel seed points (Figure 2b) are found by locating intensity minima between paired 1-D edges found within a given distance of the minima. These positions are refined and verified by testing for the existence of a pair of sufficiently strong 2-D anti-parallel (opposite direction) edges, using a set of directional kernels [34] in a circular area around a seed.

2. **Iterative tracing**: Starting from a seed point, the algorithm iteratively (Figure 2c) traces the vasculature by repeatedly detecting anti-parallel edges forming a blood vessel boundary, placing a blood vessel centerline point midway between the edge locations (vessel centerlines), and stepping along the associated direction (Figure 2c) . Landmark locations are determined from the meeting and crossing of these centerline traces (Figure 2d).

3. **Landmark refinement** involves the use of a separate mathematical model for the bifurcation and crossover locations to refine their locations to sub-pixel accuracy by fitting lines to the traces near intersections and then finding the closest point to these lines [35].

2.2 Spatial Mapping and Referencing by Quasi-Invariant Indexing

Following feature extraction, the main spatial referencing algorithm [31, 32] is applied. This algorithm is based on the idea of quasi-invariant feature indexing [22, 28, 36]. The following subsections summarize this method, describing it in two parts: an off-line spatial mapping phase, and an on-line spatial referencing phase.

2.2.1 Off-line spatial mapping

This must occur prior to surgery, from diagnostic (off-line) images of the retina that is to be treated. The goal of this phase is to generate a complete spatial map of the retina that can be used as a reference coordinate system. An associated goal is to pre-compute as much information as possible to make the on-line spatial referencing fast. The diagnostic images are traced to extract vascular landmarks and centerlines. The images are registered using the hierarchical image registration and mosaicing algorithms of Can et al. [8, 7]. This results in accurate geometric transformations linking each pair of images. Then, in each image, pairs and triples of landmarks that are reasonably close to each other (within about 20% of the image width) are identified and formed into "landmark constellations". A vector of similarity invariants — geometric measurements that do not change under the application of a similarity transformation (scale, rotation, and translation) — is computed from the locations and orientations of the landmarks in each constellation (Figure 3). These measures are in fact quasi-invariant [4] because the similarity transformation is only approximate, and therefore the vectors are called quasi-invariant feature vectors (QIFVs). QIFVs computed from all constellations in all diagnostic images are stored in the spatial map using k-d trees, as in [3], for fast lookup during the on-line phase. So, the complete spatial map consists of a set of images, their traces, Euclidean distance maps of these traces, a set of pair-wise 12-parameter quadratic spatial transformations linking the images, and the k-d tree indexing database.

2.2.2 On-line spatial referencing

During surgery, image frames are captured periodically (e.g. 30 frames/sec) from a digital camera. Each of these frames must be registered to the pre-computed spatial map. For each image, I_s, the following procedure is applied. I_s is traced to extract vascular landmarks and centerlines, and the landmarks are grouped into constellations, as above. For each constellation, the quasi-invariant feature vector (QIFV) is computed (Figure 3), and the k-d tree is searched for the closest matches. Each such match determines a correspondence for two or three landmarks (depending on whether size two or size three constellations are used) between the on-line image, I_s, and one of the off-line diagnostic images. The correspondences in each match are used to generate an initial similarity transformation. This in turn is used to select the (approximate) nearest diagnostic image, I_j, to the on-line image and generate a similarity transformation between them. This new transformation is tested for accuracy (verification) and refined (Figure 5) in a series of steps that lead to an image-wide, 12-parameter transformation. At any point during this verification and refinement, if the alignment of I_s and I_j is too inaccurate, the hypothesis is rejected and another initial match is considered. If the transformation between I_s and I_j is verified image-wide, then spatial referencing succeeds because the relationship between I_j and the entire spatial map is computed in advance. Experiments show that the median number of matches tested is 2.

The verification and refinement steps both depend on generating correspondences between vascular centerline points in I_s and I_j. Each centerline point, $\mathbf{x}_i$, from the on-line image, I_s, is transformed onto the off-line image diagnostic image, I_j. From the transformed location, $\mathbf{x}'_i$, the closest centerline point, $\mathbf{u}_i$, is found quickly using a pre-computed digital distance map [5] (Figure 4). For verification, the median of the distances between all pairs of points $\mathbf{x}'_i$ and $\mathbf{u}_i$ is computed and compared to an alignment accuracy threshold (usually about 1.5 pixels). For refinement, the pair of points $\mathbf{x}_i$ and $\mathbf{u}_i$ becomes a (temporary) match. Using $\mathbf{u}_i$ and the local curve normal, $\boldsymbol{\eta}_i$, as a local, linear approximation to the centerline contour, the centerline correspondence generated is effectively a point-to-line correspondence. These are used to refine the transformation (Figure 5) by minimizing the robust error norm given by:

$$E(\mathbf{p}) = \sum_i \rho\big([\mathbf{u}_i - \mathbf{M}(\mathbf{x}_i; \mathbf{p})] \cdot \boldsymbol{\eta}_i \;/\; \sigma\big), \tag{1}$$

where $\mathbf{p}$ is the set of parameters determining the quadratic, affine, or similarity transformations, represented generically as $\mathbf{M}(\cdot;\cdot)$ (see [8, 31]). Here $\rho(\cdot)$ is the sub-quadratic loss function of an M-estimator [14, 19, 33], and σ is a robustly estimated error standard deviation. Statistical robustness properties of the M-estimator are used to guard against spurious traces in the on-line image and missing traces in the diagnostic image in the spatial map. Typically, minimization of (1) uses the iteratively re-weighted least-squares algorithm [15], with only a few iterations. Convergence of the above alternating matching and minimization procedure takes only a few iterations.

The refinement and verification steps are applied first locally in a small region around the constellation using an affine transformation and then globally using an affine transformation and again using the 12-parameter quadratic transformation.

3 Minimizing the Cost of Spatial Referencing

The fundamental goal of this paper is to develop methods to maximize the speed of spatial referencing. To this end, it is important to realize that spatial referencing is considered complete when an accurate 12-parameter, image-wide transformation has been estimated and verified. Thus, speed can be maximized if computations not leading to this outcome are eliminated.

Feature extraction (tracing) is the most expensive computation, consuming roughly 75% of the total time, despite the significant computational reduction resulting using an exploratory algorithm [6]. Our goal therefore is find ways to minimize the amount of feature extraction computation for each image frame. The feature extraction computation provides landmark locations for indexing and centerline points for verification and refinement. In examining ways to improve the efficiency we consider first the relationship between feature extraction on the one hand and refinement and verification on the other. Changes made to improve efficiency here will have a significant impact on redesigning the landmark extraction part of the tracing computation.

3.1 Minimizing the Feature Extraction Required for Verification and Refinement

The features used in verification and refinement are formed by selecting a subset of the centerline trace points found in the on-line image, I_j. The initially tested sampling rate is 1/4 the points [29]. Sampling works for two main reasons. First, there are far more points in a complete trace than are needed. Typical images contain about 4000 on-line trace points. Second, since blood vessel orientation between landmarks changes slowly, a complete sampling is redundant and a relatively coarse sampling provides a sufficient set of constraints.

We can take this sampling idea one simple step further and use the seed points detected during grid analysis as the sampling of the vascular structure (Figure 2). Since the grid lines are regularly spaced, the seed points provide a regular sampling throughout the image. Experimentally, we have found that replacing the sampling of the traced vasculature with the seed points has no effect whatsoever on the success rate of our spatial referencing algorithm and the average alignment error changes from 0.65 pixels to 0.81 pixels. On average, on-line grid analysis produces 167 seed points, all used in verification and refinement.

Computationally, the implications of using sampled seeds in verification and refinement extend far beyond the improvement in efficiency in the verification and refinement steps. Beyond seed point detection and verification, the job of the tracing algorithm in support of spatial referencing is reduced to find constellations of (two or three) landmarks. Previously, centerline points were needed in addition to landmarks. Moreover, if the first landmark constellation found has a corresponding constellation in the spatial map and if this correspondence is found through indexing the k-d tree, then it will likely lead to a verified global transformation. Thus, the required job of tracing the vasculature — the most time-consuming phase of the spatial referencing computation — can be finding constellations of landmarks. In the best case only one constellation is needed and on average only a few. In short, it is possible to completely eliminate the need for further tracing after good constellations have been generated.

3.2 Minimizing Tracing Computation for Landmark Constellations

The next innovation is to substantially reduce the computations in locating landmark constellations. The idea here is to use the grid analysis to *predict* the locations of landmarks and constellations and then *prioritize* the selection of seed points for tracing based on these predications. A first step toward this was taken by Shen et al. [30] in which the prioritization was based on landmarks alone. Here we extend this idea to prioritization based on constellations. We will also introduce a hierarchical grid analysis technique to further improve efficiency.

3.2.1 Prioritizing for Early Landmark Extraction

The seed points detected in the first step of the exploratory algorithm (Section 2.1) give initial locations and orientations of blood vessel centerline points. The grid lines on which the seeds are detected partition the image into boxes. The distribution of the orientations of the seed points over a grid box provides strong hints about the presence or absence of a crossing/branching point within the grid box (Figure 6). Boxes where the seed points have approximately the same orientation are likely to just be intersecting a single blood vessel or two parallel vessels (Figure 6(e) and (f)), and are therefore not likely to contain a landmark. On the other hand, boxes containing a diversity of seed point orientations are much more likely to contain landmarks (Figure 6(b)-(d), (g)). Tracing should be applied to seed points in these boxes first.

Shen et al. [30] proposed a method for placing a priority on each box by combining a measure of the seed orientation diversity (illustrated by the examples) with a measure of the strength (contrast) of a seed point's vessel boundaries. Let N be the number of seed points on a grid box b, let s_i denote the estimated strength (contrast) of each seed point, and let θ_i be the estimated blood vessel tangent direction at each seed. Then the "strength-weighted angular diversity" measure of grid box b is given by:

$$q(b) = \sum_{i=1}^{N} s_i - \left\| \sum_{i=1}^{N} s_i [\cos 2\theta_i, \sin 2\theta_i] \right\| + N. \tag{2}$$

The angle values are doubled so that for a box containing a pair of seeds with exactly opposing orientations (anti-parallel) and the same strengths, the first two terms will exactly cancel. By contrast for a box having two equally strong perpendicular seeds, the second term will be 0, producing a much higher value. We have augmented the original equation from [30] by adding the N term at the end to bias $q(b)$ toward boxes with more seed points.

Shen et al. [30] proposed prioritizing tracing by the value of $q(b)$ (see Section 3.2.3 for details). This strategy was shown to result in nearly optimal schedules for extracting landmarks, as measured by the number of landmarks extracted as a function of the number of tracing steps. One limitation of this idea in the context of our desire to perform indexing with local constellations of landmarks is that the first few landmarks thus detected may be scattered throughout the image, and not close enough to form a constellation. In other words, successful spatial referencing really requires prioritization that results in the best-possible local constellations of landmarks, and not just individual landmarks.

3.2.2 Prioritizing for Early Constellation Extraction

Recall that the number of landmarks in a constellation is k, which is either 2 or 3, depending on which QIFV is used. The distance between landmarks in a constellation must be at most R (Section 2.2.1) and at least r, which is empirically determined to be about 2% of the image width. The lower bound eliminates instabilities in constellations formed from landmarks that are too close to each other.

Intuitively, once a landmark is formed during tracing, the tracing computation should be focused on other nearby boxes containing high potential for landmarks. Here, "nearby" means distance between r and R. When looking for constellations of three landmarks, and a second landmark has been found, potential landmark locations that are nearby both landmarks should have the highest priority. This is illustrated in Figure 7.

This idea is easily quantified. During tracing, let $n(b)$ be the number of landmark locations, $\mathbf{q}$, that have already detected and are within distance R but no closer than r of the center of grid box b. Larger values of $n(b)$ means that it is more likely that a landmark found in b will yield a constellation. Therefore, we can augment the earlier strength-weighted angular diversity measure of (2) to give a new, combined priority measure for a grid box,

$$Q(b) = (n(b) + 1) \cdot q(b). \tag{3}$$

3.2.3 Prioritized Tracing Computation

Here's an outline of the prioritized tracing computation. Following grid analysis and seed point detection, $q(b)$ is computed for each grid box, $n(b)$ is initialized to 0, and $Q(b)$ is initialized to $q(b)$. The boxes are then placed in a priority queue. The following steps are executed until the priority queue is empty:

1. The highest priority box, b, is removed from the priority queue. Tracing is run for all seed points on the boundary of b, with tracing directed toward the box interior.
2. The endpoints of traces (within b or on the boundary) are tested to see if they form intersections. Any such intersections are formed into landmarks.
3. For each new landmark location, $\mathbf{q}$, constellations are formed with nearby landmarks, if possible.
4. For each new landmark location, $\mathbf{q}$, $n(b')$ is updated for nearby boxes, b'. $Q(b')$ is recalculated for these boxes, and the box positions in the priority queue are adjusted accordingly.
5. When a trace stops on the boundary of box b at a point that is not a previously detected seed, a new seed is added for each neighboring box b' (usually just one, but as many as three at a corner) on which the new seed resides. In this case, $q(b')$ and, as a result, $Q(b')$ are recalculated, and the position of b' is adjusted in the priority queue. If b' has already been

traced, then it must be returned to the priority queue. When it is brought up for tracing again, only the untraced seed points will be traced.

These steps are illustrated in Figure 8.

3.3 Integrating Tracing And Spatial Referencing Computation

Integration of the prioritized tracing procedure with spatial referencing is straightforward:

Grid analysis and seed point detection: The grid analysis and seed point detection algorithms are applied to produce a set of seed points $\mathcal{S}$ and the initial values of $n(b)$, $q(b)$ and $Q(b)$ are calculated as just described.

Pre-emptive tracing: The prioritized tracing algorithm is run, but it is pre-empted (temporarily halted) whenever a landmark constellation is found.

Indexing: For each constellation, the quasi-invariant feature vector (QIFV) is computed and the two nearest constellation matches are found through indexing into the k-d tree. These matches must be for different constellations found in the off-line images (not the same constellation in different off-line images).

Refinement and verification: Each constellation matches generates an associated set of landmark correspondences. These correspondences are used to generate an initial transformation between the on-line (surgical) image and one of the off-line (diagnostic) images in the spatial map. This transformation is refined and verified as described in Section 2.2.2. The set of seed points, $\mathcal{S}$, together with a sampling of the current trace points is used as the sampling of the vascular centerline points from the on-line image.

Termination: The algorithm cycles through pre-emptive tracing, indexing, and refinement and verification until one constellation match leads to globally verified transformation or until the priority queue of boxes is exhausted.

If the priority queue is exhausted then the next three best constellation matches are tried for each constellation, as consistent with the original implementation. The number of landmarks in a constellation is set at either two or three. Our experiments use two.

The power of the integration of tracing and spatial referencing can be illustrated by considering the ideal case: the prioritized tracing of grid boxes produces a constellation of two landmarks from its first two boxes, and these lead to a verified spatial referencing result. In this case, just 2 out of 900 boxes (for a 30×30 grid) are traced! Such an example is shown in Figure 10(a).

3.4 Hierarchical Grid Analysis

A hierarchical component can be added to the integrated tracing and spatial referencing algorithm to further improve its computational performance. The motivation is that large blocks of the image where there are few blood vessels, especially on the periphery of the retina, can be identified and hence ignored based only on a coarse analysis. Similarly, coarse analysis can be used to identify and focus computation on feature-rich regions.

Several techniques are needed to realize this idea. The image is initially divided up into $m \times n$ grid boxes formed by evenly spaced horizontal and vertical lines. These grid boxes form the coarsest level of the hierarchy, L_0 and as a consequence, m and n are typically small (e.g. $m = n = 4$). Grid boxes are formed at subsequent levels of the hierarchy, $L_i, i = 1, \ldots, m$, are formed recursively by subdividing a subset of the boxes at level $i - 1$ (see Figure 6), based on priority. Level L_m is the finest (full) resolution, and corresponds roughly to the original boxes in the non-hierarchical version of the algorithm. Tracing is only applied at level L_m. To do this, $m + 1$ priority queues, P_i, are formed, one for each level. Priorities for boxes at levels 0 through $m - 1$ are based on the strength-weighted edge-angle diversity measure (Equation 2), $q(b)$. Priorities at level L_m are based on $Q(b)$ (Equation 3), which combines $q(b)$ with $n(b)$, the number of landmarks that have already been detected near the box.

Initially, the $m \times n$ grids are analyzed to detect seed points as described in Section 2.1. The set S is initialized to contain these seeds. Then, each level 0 box, b, is analyzed to compute $q(b)$, and the boxes are placed in priority queue P_0 according to $q(b)$. The priority queues at the other levels are initially empty. The following procedure is repeated until one transformation has been globally verified or until all queues are empty (and no transformation is estimated):

1. For $i \in [0, \ldots, m - 1]$,

    (a) Remove the top 2^i boxes from priority queue P_i. This takes only one box from level L_0

and at most half of the boxes at any other level.

(b) Subdivide each removed level i box, b_r, to create four level $i+1$ sub-boxes by drawing a vertical and a horizontal line through the center of the box, and doing seed detection along the new lines. Add the seeds to the appropriate boxes as well as to the overall set of seeds, S. Distribute seeds from the boundary of b_r to the new sub-boxes.

(c) Calculate the priority in each sub-box and enter into priority queue P_{i+1} for level L_{i+1}. At all levels except L_m use $q(b)$ to prioritize the boxes; at level L_M use $Q(b)$.

2. Using priority queue P_m at level L_m, apply the combined prioritized tracing and spatial referencing algorithm described in Section 3.3. Stop with sucess when an image-wide, verified transformation is found. Otherwise, temporarily halt when 2^m boxes have been traced and repeat the above.

Several features of this procedure require further explanation. First, only a small subset of the boxes at each level is chosen for expansion based on the idea that computation should be focused only on the highest priority regions. Second, if a failure occurs within the first few (2^m) boxes at the finest level, the procedure returns to the coarsest level in order to refocus the computation elsewhere in the image. Third, when a trace hits the boundary of a box where there previously was not a seed, a seed must be added (see Section 3.2.3). This must be propagated up the hierarchy. Note, however, that only at the finest level, L_m, is a box ever re-entered in the priority queue. Overall, this hierarchical algorithm saves an additional 20% of the computation time over the non-hierarchical approach, which is significant for this application.

4 Experimental Analysis

The proposed algorithm for integrated tracing and spatial referencing was evaluated on 373 1024×1024 images captured from 18 healthy retinas using a TOPCON TRC/IMAGEMAP fundus camera system. Our test methodology on the 18 data sets is based on the "leave-one-out" approach. Suppose for a given data set (single eye) there are M images, $\mathcal{T} = \{I_1, \ldots, I_M\}$. One separate test is run for each image I_j. The image set $\mathcal{T} - \{I_j\}$ 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 to which the integrated algorithm is applied. Performance statistics are gathered by applying the algorithm to all different values of j and to all data sets.

Figure 10 provides several example results of the procedure for integrating spatial referencing with prioritized tracing. The illustrations show varying levels of success. Figure 10a shows the ideal case in which spatial referencing succeeded after tracing just 2 boxes. This represents less than 1% of the image. Figure 10b is an average case that required 7 boxes to be traced. Figure 10c shows a less common case that required 55 boxes to be traced. Figure 10d illustrates a complete failure where 133 boxes were traced, and yet spatial referencing was not successful. The failure is due to having little overlap with the spatial map.

The significance of scheduling grid boxes based on landmark constellation prioritization (LC — using $Q(b)$ as the measure) can be measured by comparing it to scheduling based on landmark prioritization (LM) alone and to random scheduling. For landmark priorization, $q(b)$ is used in place of $Q(b)$ in the algorithm of Section 3.2.3; for random scheduling, the boxes are picked in random order. In all cases, whenever a landmark is detected, the algorithm checks for the formation of a constellation, and whenever a constellation is formed, indexing and spatial referencing are applied, all as described in Section 3.3. The results are shown in Table 1 and Figure 11 for a 2.2GHz Pentium Xeon processor. Clearly, random scheduling requires more tracing, and overall computation before spatial referencing succeeds. Landmark (LM) based prioritization reduced the tracing computation to a large degree — by a factor of 2.5 on the number of trace points and 4.5 in the number of traced boxes when compared to random prioritization. Landmark constellation (LC) prioritization further improved the efficiency of spatial referencing, but only modestly. This is mainly due to the fact that a large distance range (tolerance) between landmarks is allowed for forming landmark constellations.

The hierarchical grid analysis improved the overall computation time by 20%. Most of these improvements were in the seed point detection time and spatial referencing time, the latter being directly attributable to the reduced number of seed points. In fact, the reduced number of seed points causes four additional failures in the spatial referencing algorithm. The number of traced boxes and traced points increases somewhat in the hierarchical algorithm because fewer boxes are available in the initial prioritization.

For the vast majority of on-line images, fewer than 400 trace points are required to complete spatial referencing for both landmark-based (LM) and constellation-based (LC) scheduling, which is only about 9% of the full-tracing effort. Unfortunately, 15 images generated more than 1000 trace points before succeeding, which explains why all the averages in Table 1 are substantially larger than the medians. Most of these images are taken of the periphery of the retina and have relatively low overlap with the other images that form the leave-one-out spatial map. Moreover, these images tend to have fewer landmarks (average of 16 vs. 39 when the images are fully traced). An important source of these extreme cases is our evaluation method itself. The leave-one-out strategy for evaluating the spatial referencing algorithm occasionally results in a map with less than complete coverage. Further evidence for this is the increased number of constellations tested of the 15 images. For incomplete spatial maps, many constellations found in peripheral images are not in the spatial map at all, which resulted in an average of 38 constellations for the 15 worst images as compared to 5.5 overall. Our future real-time tests will use on-line images taken during surgery, and will construct the spatial map using all diagnostic images. Ensuring that the spatial map is complete should reduce the effects of peripherel images substantially. In general, however, finding ways to reduce the time for the worst cases will be important for ensuring predictable real-time performance, and therefore is an important consideration in ongoing research.

Finally, an important parameter in the algorithm is the number matching constellations extracted from the k-d tree for each landmark constellation. In our earlier implementation, where all constellations were extracted before spatial referencing [29, 31], the five nearest neighbors were extracted. The nearest was tested first for each constellation. If all failed, then the second nearest was tested for each constellation, etc. In the integrated technique, the two best are tested for each constellation before tracing is allowed to continue. If spatial referencing fails for the best two for every constellation, then after all tracing is completed, the third best is tried for each constellation, then the fourth best for each and then the fifth. Increasing the number tried before tracing continues tends to improve the performance for the worst images, but also increases the overall average and median times. This trade-off will require further investigation when true on-line image sequences are available.

5 Discussion and Conclusions

This paper has described an adaptive method to minimize the amount of feature extraction computation to achieve spatial referencing. The first innovation was to use tracing seed points from the fast, grid-based analysis of the image as a sampling of the retinal vasculature in verifying and refining transformations generated through invariant indexing. This implies that the actual exploratory tracing algorithm needs only to extract landmark and landmark constellations sufficient for invariant indexing to generate initial transformation estimates. Therefore, the second innovation, building upon the work of Shen et al. [30], was to prioritize tracing to extract landmarks and landmark constellations as quickly as possible. The third innovation was to embed the grid analysis and tracing computations in a hierarchial framework that focused computations on only the most promising regions. These techniques were integrated into spatial referencing to aggressive test each landmark constellation as a start to the registration process as soon as it was detected. The end result is a remarkably fast, highly adaptive and effective method to achieve spatial referencing.

This work has demonstrated several design principles. First, it has shown the power of exploratory image analysis. Algorithms based on significant processing of each pixel would be several orders of magnitude slower than the ones described here. Second, it has resulted in improved methods to prioritize spatially the already sparse computations associated with exploratory tracing of vasculature. Third, it has demonstrated that feature extraction can be further minimized by an effective tight integration of on-line vascular tracing with the spatial referencing of retinal images in an opportunistic manner. These design principles and many of the techniques themselves can and will be applied to new versions of the spatial referencing registration algorithm.

Interestingly, all of this was achieved on conventional off-the-shelf (COTS) computer hardware. The execution times on a generic 2.2GHz Pentium Xeon processor already achieve a speed of roughly 20 frames/sec on 1024×1024 size images. As computing speeds inevitably increase, the achievable frame rates will increase accordingly.

Within the broader context of the intended application, much work remains. For example, the proposed methods can be applied based on regions of surgical interest. The position of the aiming laser, often present in instruments, can serve as another clue for prioritization. Reliability and predictability are key issues in the design of hard real time and deadline-driven spatial referencing systems. Finally, parallel computing techniques may be adopted to speed up the whole tracing/spatial referencing procedure. In this context, the process of tracing the image in small grid boxes enables simple coarse-grain parallelization by image tiling.

6 Acknowledgements

The authors would like to thank the staff at the Center for Sight, especially photographers Gary Howe and Mark Fish, for extensive help with image acquisition. We are thankful to Dr. Ali Can for discussions and valuable suggestions. Various portions of this research was supported by the National Science Foundation Experimental Partnerships grant EIA-0000417, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821), the National Institutes for Health grant RR14038, and by Rensselaer Polytechnic Institute

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

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

[6] 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.

[7] 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.

[8] 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.

[9] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3):263–269, September 1989.

[10] J. L. Coatrieux, M. Garreau, R. Collorec, and C. Roux. Computer vision approaches for three-dimensional reconstruction: Review and prospects. *Critical Reviews in Biomed. Eng.*, 22(1):1–38, 1994.

[11] A. R. Cohen, B. Roysam, and J. N. Turner. Automated tracing and volume measurements of neurons from 3-d confocal fluorescence microscopy data. *Journal of microscopy*, 173(2), February 1994.

[12] H. Davson, editor. *The Eye*, volume 1–4. Academic Press, New York, 1962.

[13] J. Federman, editor. *Retina and Vitreous*. The C.V. Mosby Company, St. Louis, 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] J. M. Krauss and C. A. Puliafito. Lasers in ophthalmology. *Lasers In Surgery And Medicine*, 17:102–159, 1995.

[17] P. A. Laplante, editor. *Real-Time Systems Design and Analysis: An Engineer's Handbook, 2nd ed.* IEEE Press, 1996.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

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

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

[24] E. D. Oberg, S. F. Barrett, and C. H. Wright. Development of an integrated automated retinal surgical laser system. *Biomed Sci Instrum*, 33:77–81, 1997.

[25] R. Polli and G. Valli. An algorithm for real-time vessel enhancement and detection. *Computer methods and programs in Biomedicine*, 52:1–22, 1997.

[26] K. Ramamritham. Predictability: Demonstrating real-time properties. *ACM Computing Surveys*, 28(4(es)), Dec 1996.

[27] K. E. Rayner, editor. *Eye Movements and Visual Cognition: Scene Perception and Reading*. Springer Series in Neuropsychology. Springer-Verlag, New York, 1992.

[28] 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.

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

[30] 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.

[31] 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 IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 79–86, 2001.

[32] 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*, accepted 2002.

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

[34] Y. Sun, R. Lucariello, and S. Chiaramida. Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary arteriograms. *IEEE Transactions on Medical Imaging*, 14:242–248, June 1995.

[35] 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.

[36] 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.

Figures

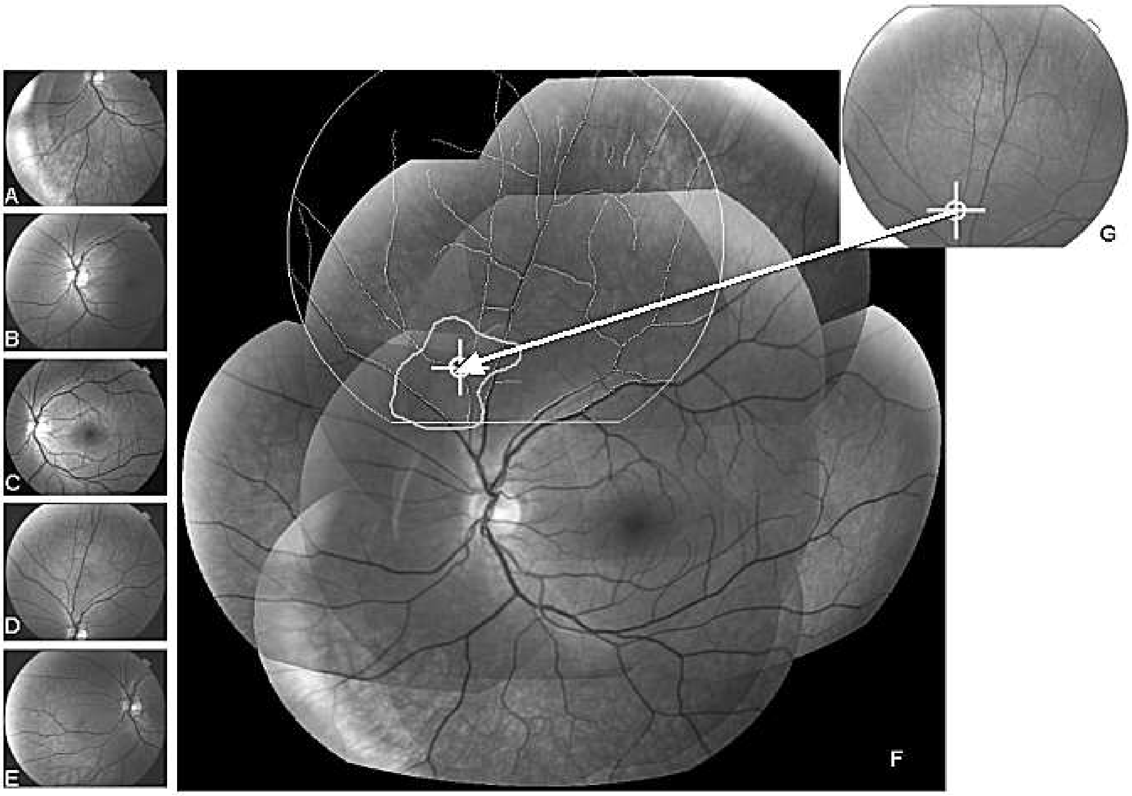

Figure 1: Illustration of spatial mapping and referencing as registration problems. Spatial mapping is performed prior to surgery. Diagnostic images (five are shown in A through E) are mosaiced together to form a wide-extent spatial map of the entire retina. The spatial map consists of the image mosaics (panel F), the surgical treatment region (green outline in Panel F) drawn by the physician, and various data structures (described in Section 2) to support fast spatial referencing during surgery. Spatial referencing is the problem of registering each image frame captured by a camera (Panel G) onto the spatial map, effectively estimating the location of a laser beam (illustrated as yellow cross-hairs on Panel G) relative to the spatial map. The red arrow illustrates this mapping. Spatial referencing can for example, determine whether or not the surgical laser is aimed within the desired treatment region.

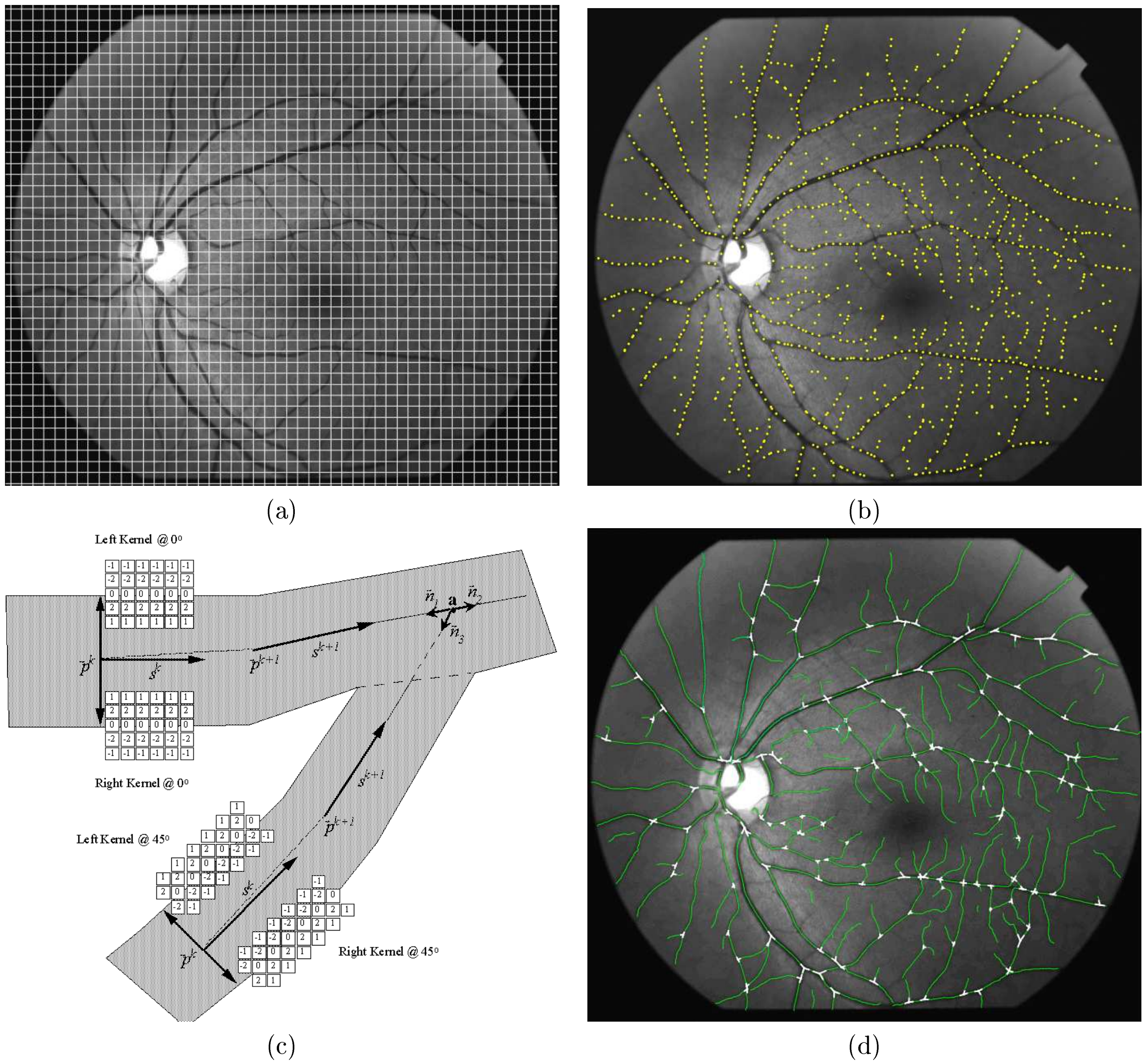

Figure 2: Illustrating the tracing algorithm for vascular feature extraction. Panel (a) shows the grid used for analyzing intensity statistics and gathering seed points. Grids of 80 vertical and 80 horizontal lines, as shown, are used for off-line computations, whereas grids of at most 30 vertical and 30 horizontal lines are used for on-line computations. Panel (b) shows the seed points detected from the grid analysis. Panel (c) illustrates the application of templates for iterative tracing. Panel (d) shows the blood vessel centerline points and landmarks (branching and cross-overs of the retinal vasculature) detected.

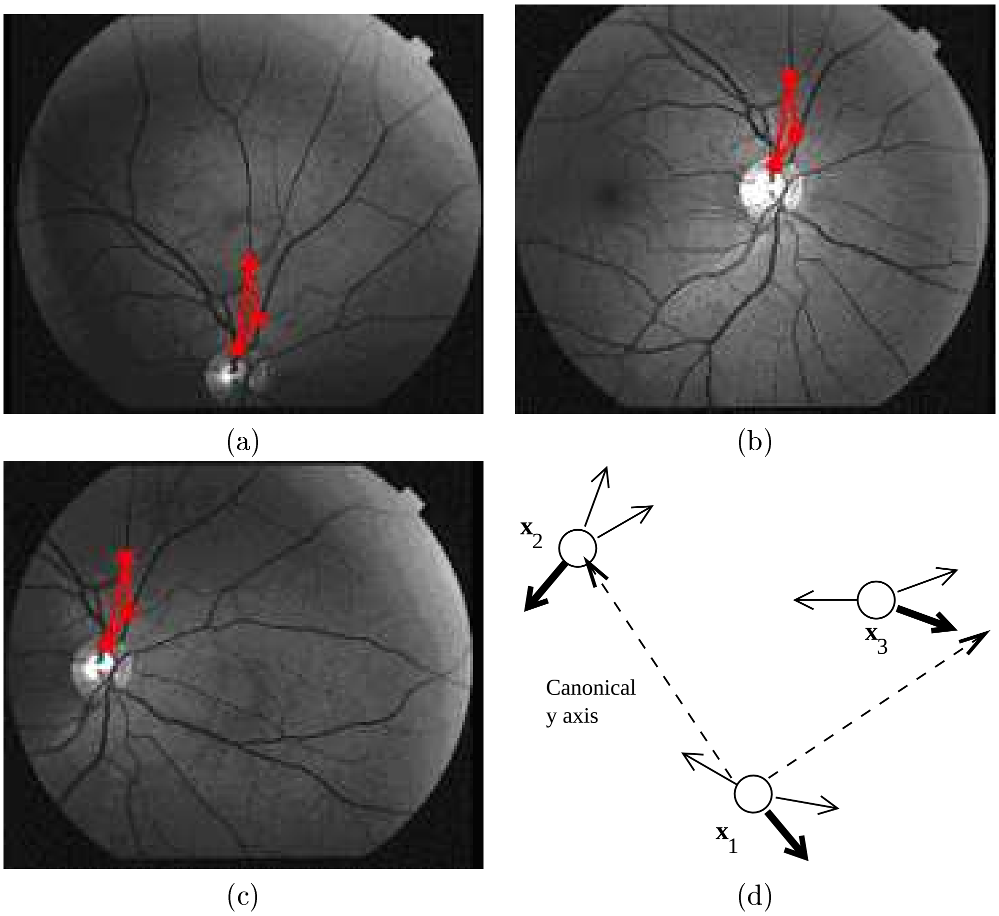

Figure 3: Illustrating the use of invariant indexing for spatial referencing. Panels (a)-(c) show three sample image frames where the same vascular landmark constellation appears in red. Even though the viewpoints are different, the relative positions of the landmarks remains approximately the same. Panel (d) illustrates the calculation of the quasi-invariant feature vector from a constellation of three landmarks. The locations of landmarks $\mathbf{x}_1$ and $\mathbf{x}_2$ are used to construct a coordinate reference frame. The coordinates (α, β) of the third landmark, $\mathbf{x}_3$, in this frame form the first two components of the QIFV. The orientations of the vascular segments are mapped into the range $(-\pi, \pi]$ in this frame, and the most negative angle (shown in bold) for each landmark forms the other three components.

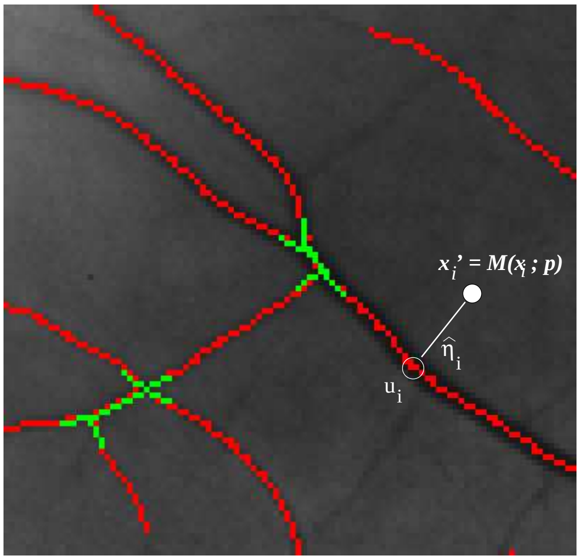

Figure 4: Establishing matches for verification and refinement. Based on the current transformation parameter estimate, point $\mathbf{x}_i$ from the on-line image is mapped to location $\mathbf{x}'_i$ in the selected off-line image. The nearest tracepoint $\mathbf{u}_i$ in the off-line image and its local curve normal $\hat{\eta}_i$ are found from the digital distance map that coincidences with the off-line image.

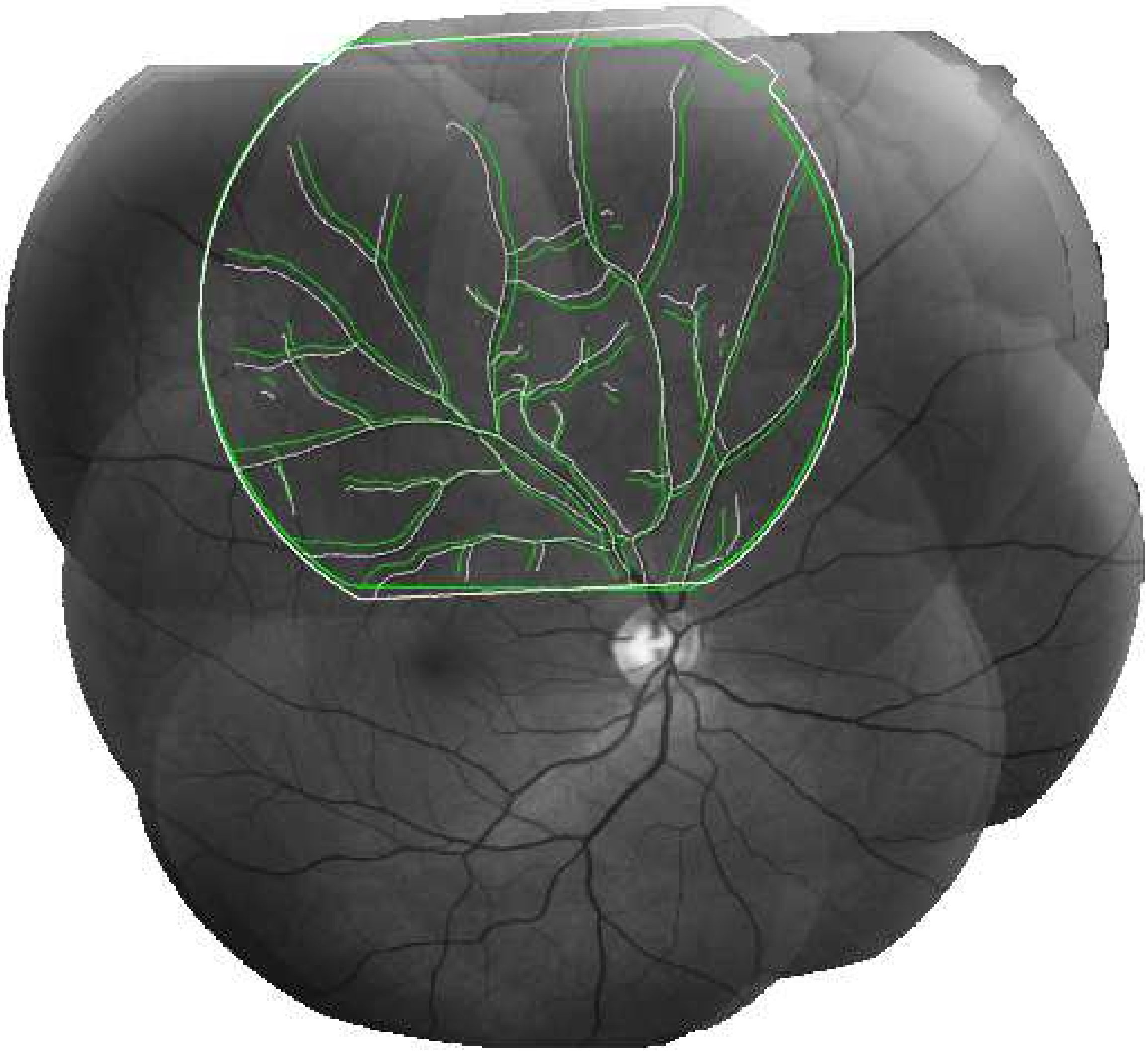

Figure 5: Illustrating the transformation refinement process. The transformation of the on-line image frame, I_1, to the spatial map is illustrated by overlaying its transformed traces on top of a mosaic of the off-line images. This shows some alignment error from local estimation that follows invariant indexing. Steps of matching between the vascular centerline points in I_s and one of the off-line diagnostic images, I_j, stored in the spatial map generates constraints to refine this transformation. Several iterations of refinement yields the final (correct) alignment shown in white. See Figure 10 for more results.

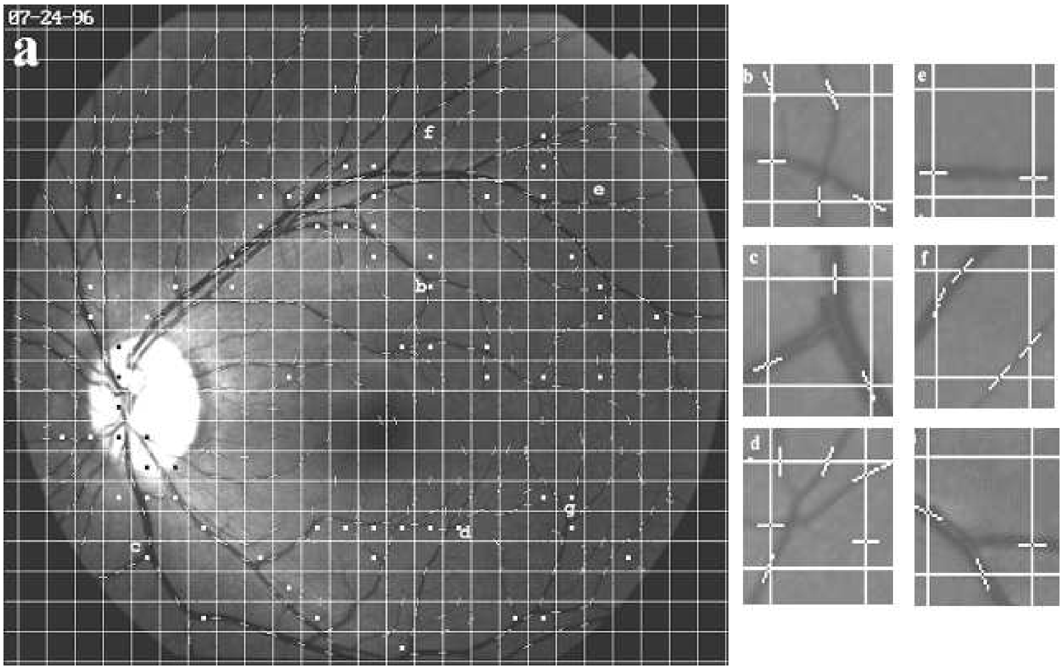

Figure 6: Illustrating the intuition behind prioritization of grid boxes for tracing. The boxes are rectangular image regions whose corners are formed by the intersection of grid lines (a). Six boxes are labeled (b)-(g) on the left and shown at high resolution on the right. Boxes (b), (c), (d) and (g) illustrate cases when the seed points provide strong clues about the presence of landmarks. In contrast, Boxes (e) and (f) illustrate cases when the seeds points suggest the absence of landmarks.

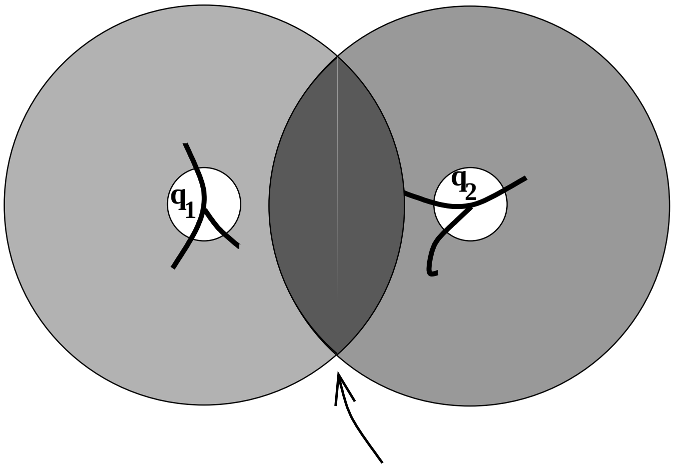

Figure 7: Grid box boxes near landmark locations should get higher priority. Two landmark locations, $\mathbf{q}_1$ and $\mathbf{q}_2$, are shown. The inner and outer circles surrounding each are radius r and R respectively and represent the minimum and maximum distances over which the landmarks can participate in a constellation. The most promising location to search for more landmarks is the dark grey region (indicated by the arrow) because a landmark there would participate in one constellation of triples and two constellations of doubles.

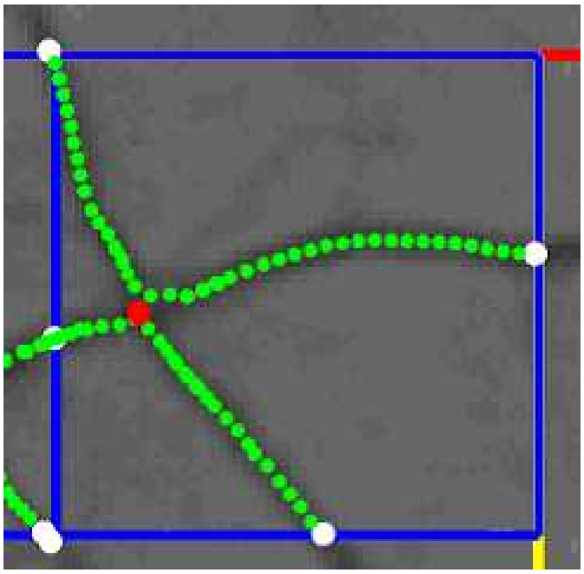

Figure 8: Illustrating tracing in a grid box. Tracing is initialized at five seed points (shown in white) on the boundary of the box. One runs out of the box and terminates immediately. The others form the green trace points and meet to form the intersection (landmark) locations shown in red.

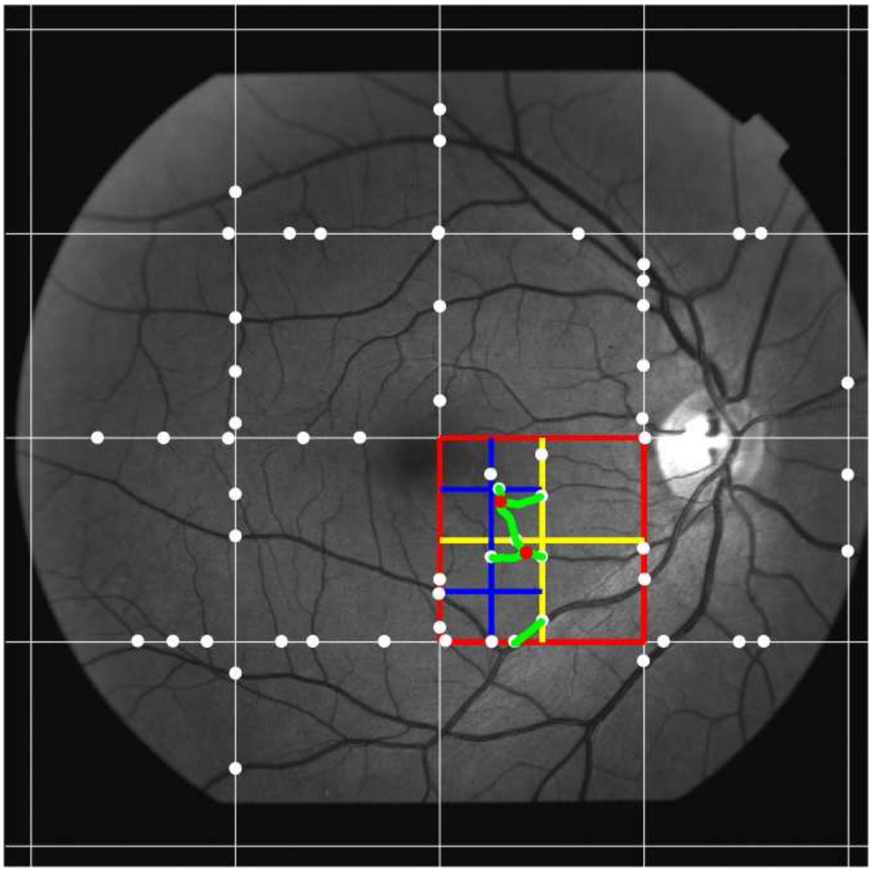

Figure 9: Illustrating the hierarchical grid analysis. The image shows the coarse resolution sampling, with several further levels of subsampling in one box. The actual trace points are shown in green. Three landmark locations are shown as red dots.

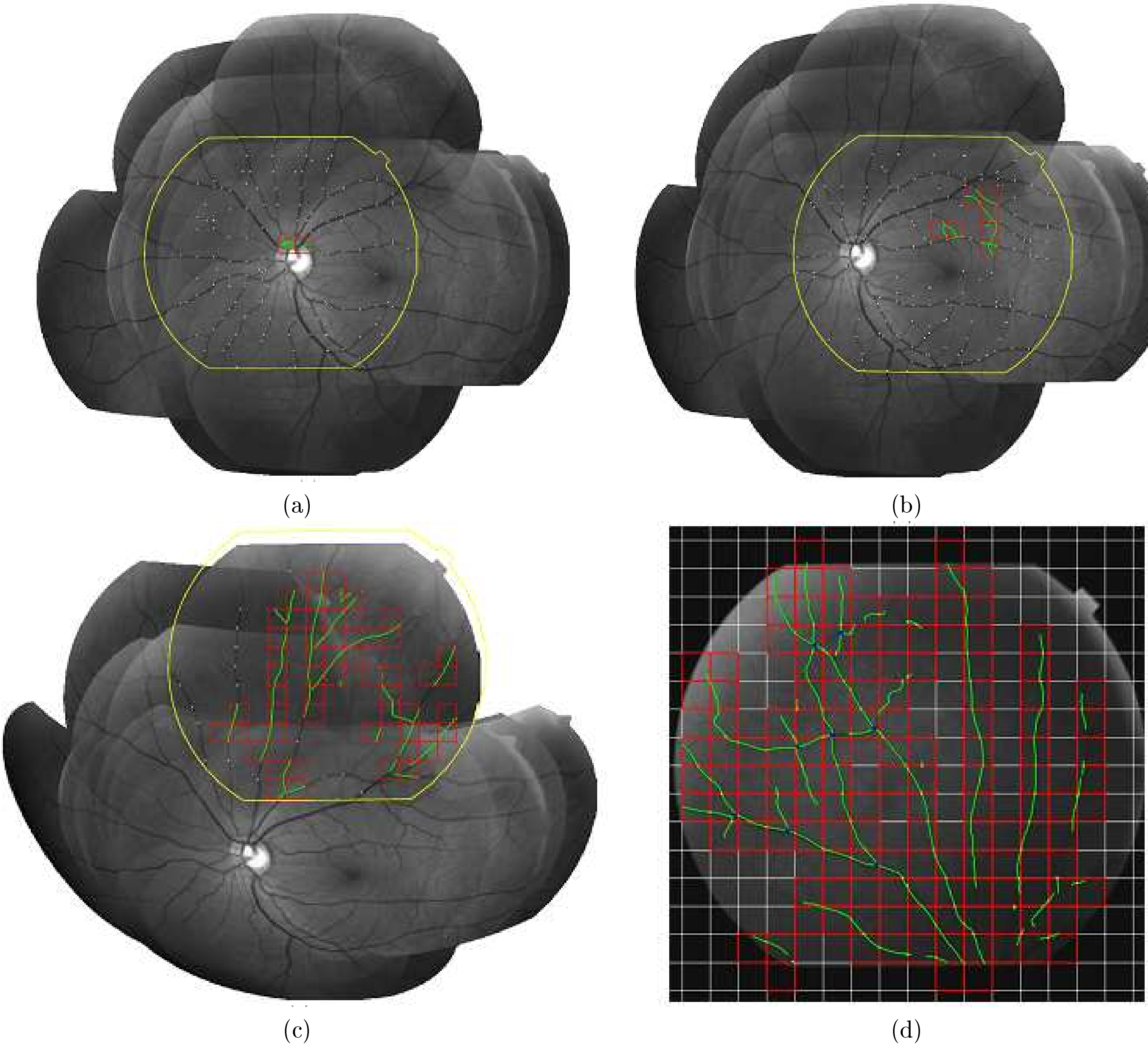

Figure 10: Sample images illustrating different levels of success. Panel (a) shows a highly successful case (140 seed points, 114 trace points, 2 boxes traced, 3 detected landmarks). Panel (b) illustrates an average case (162 seed points, 336 trace points, 7 boxes traced, 4 detected landmarks). Panel (c) illustrates a less favorable and less frequent case that typically occurs near the retinal periphery where the vasculature is sparse (84 seed points, 1445 trace points, 55 boxes traced, 5 detected landmarks). Panel (d) illustrates a case of complete failure, due to very little overlap with the spatial map (95 seed points, 2983 trace points, 133 boxes traced, 8 detected landmarks).

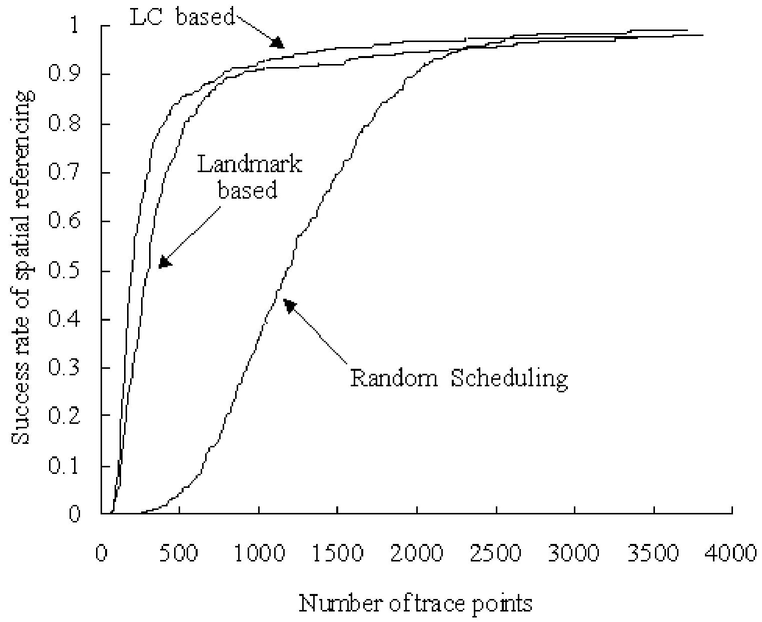

Figure 11: Illustrating the relative performance of three different scheduling algorithms over a sequence of 373 image frames (1024×1024 pixels each). The plot shows the fraction of images for which spatial referencing has succeeded for the given number of trace points. "LC based" is prioritization of tracing using prediction of landmark constellations ($Q(b)$). "Landmark based" is priorization of tracing using prediction of landmarks alone ($q(b)$). "Random scheduling" is a random ordering of boxes.

Tables

Method	Priority Levels	Avg # Traced Boxes	Avg # Traced Points	Success Rate(%)	Time (milliseconds)							
					Seeds		Tracing		Referencing		Total	
					M	A	M	A	M	A	M	A
Random	1	35.8	1047	95.7	23	22.7	22	24.4	11	15.1	55	61.5
LM	1	8.9	429	96.8	23	22.7	8	10.4	9	14.2	40	47.5
LC	1	7.1	340	96.8	23	22.7	6	8.3	9	13.9	38.3	45.5
LC	3	9.4	392	95.7	10	9.5	7.7	9.8	6	10.9	24.3	31.2

Table 1: Performance of prioritized tracing/spatial referencing based on three different scheduling schemes. All tests were done on a 2.2GHz Pentium Xeon processor. In the table, 'M' means median and 'A' means average. "Random" refers to random ordering of the grid boxes, "LM" refers to prioritizing the grid boxes based on landmarks alone, "LC" refers to prioritizing the grid boxes based on extraction of landmark constellations. The hierarchical grid analysis used $m = 2$, for 3 levels overall. All tests were on 1024×1024 pixel images.