Image Registration Using Log-Polar Mappings for Recovery of Large-Scale Similarity and Projective Transformations

Siavash Zokai and George Wolberg, *Senior Member, IEEE*

***Abstract*—This paper describes a novel technique to recover large similarity transformations (rotation/scale/translation) and moderate perspective deformations among image pairs. We introduce a hybrid algorithm that features log-polar mappings and nonlinear least squares optimization. The use of log-polar techniques in the spatial domain is introduced as a preprocessing module to recover large scale changes (e.g., at least four-fold) and arbitrary rotations. Although log-polar techniques are used in the Fourier–Mellin transform to accommodate rotation and scale in the frequency domain, its use in registering images subjected to very large scale changes has not yet been exploited in the spatial domain. In this paper, we demonstrate the superior performance of the log-polar transform in featureless image registration in the spatial domain. We achieve subpixel accuracy through the use of nonlinear least squares optimization. The registration process yields the eight parameters of the perspective transformation that best aligns the two input images. Extensive testing was performed on uncalibrated real images and an array of 10,000 image pairs with known transformations derived from the Corel Stock Photo Library of royalty-free photographic images.**

***Index Terms*—Image registration, Levenberg–Marquardt nonlinear least-squares optimization, log-polar transform, perspective transformation, similarity transformation.**

I. Introduction

DIGITAL image registration is a branch of computer vision that deals with the geometric alignment of a set of images. The set may consist of two or more digital images taken of a single scene at different times, from different sensors, or from different viewpoints. A large body of research has been drawn to this area due to its importance in remote sensing, medical imaging, computer graphics, and computer vision. Despite comprehensive research spanning over thirty years, robust techniques to register images in the presence of large deformations remains elusive. Most techniques fail unless the input images are misaligned by moderate deformations.

The goal of registration is to establish geometric correspondence between the images so that they may be transformed, compared, and analyzed in a common reference frame. Registration is often necessary for 1) integrating information taken from different sensors (i.e., multisensor data fusion), 2) finding changes in images taken at different times or under different conditions, 3) inferring three-dimensional (3-D) information from images in which either the camera or the objects in the scene have moved, and 4) for model-based object recognition. The most common task associated with image registration is the generation of large panoramic images for viewing and analysis. Image mosaics, created by warping and blending together several overlapping images, are central to this process. Other common registration tasks include producing super-resolution images from multiple images of the same scene, change detection, motion stabilization, topographic mapping, and multisensor image fusion.

This work attempts to register two images using one global perspective transformation even in the presence of arbitrary rotation angles and large scale changes (up to $5\times$ zoom). Our work is motivated by the problem of registering airborne images. These images are taken at vastly different times, altitudes, and directions. Therefore, the images differ by large rotation and scale. Also, the pitch and roll introduces moderate perspective.

In general, images of a 3-D scene do not differ by just one perspective transformation because the depth between the camera and the objects introduces parallax. A global transformation cannot align all features in such cases. We must, therefore, place constraints on camera motion and/or our 3-D scene to produce images that are free of parallax. One constraint requires the camera motion to be limited to rotation, pan, tilt, and zoom about a fixed point, e.g, on a tripod. If this constraint is not satisfied, then we may still have images free of parallax if the object's 3-D points (x, y, z) in the scene are far away from the camera, i.e., $x, y \ll z$. This means that the scene is flat and we are looking at a planar object. In either case, we assume that the scene is static and the lighting is fixed between images. Nevertheless, we have relaxed these conditions to accommodate local disparity and linear changes in illumination.

A survey by Brown [1] introduces a framework in which all registration techniques can be understood. The framework consists of the feature space, similarity measure, search space, and search strategy. The *feature space* extracts the information in the images that will be used for matching. The *search space* is the class of transformations, or deformation models, that is capable of aligning the images. The *search strategy* decides how to choose the next transformation from this space, to be tested in the search for the optimal transformation. The *similarity measure* determines the relative merit for each test. Search continues according to the search strategy until a transformation is found

Manuscript received April 27, 2004; revised October 11, 2004. This work was supported in part by an ONR HBCU/MI Research and Education Program Grant (N000140310511) and a PSC-CUNY Grant. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Luca Lucchese.

S. Zokai is with Brainstorm Technology LLC, New York, NY 10011 USA.

G. Wolberg is with the Department of Computer Science, City College of New York, New York, NY 10031 USA (e-mail: wolberg@cs.ccny.cuny.edu).

Digital Object Identifier 10.1109/TIP.2005.854501

1057-7149/$20.00 © 2005 IEEE

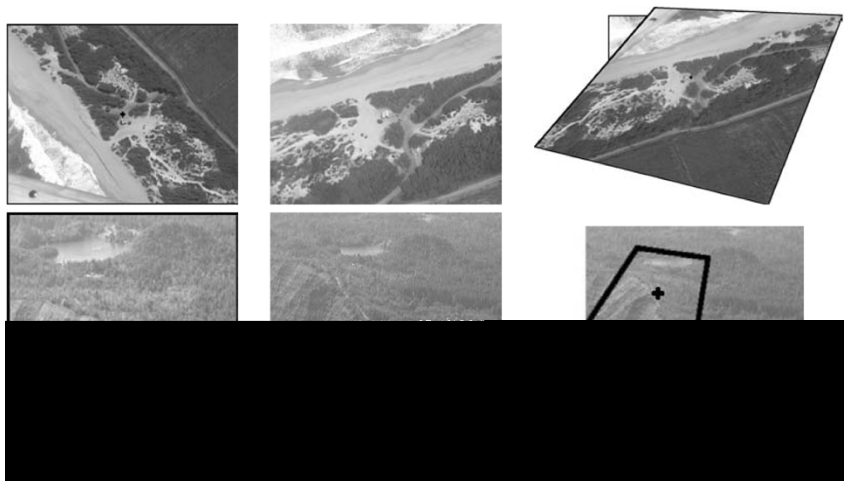

Fig. 1. Airborne imagery. (a) Observed images. (b) Reference image. (c) Registration overlays.

whose similarity measure is satisfactory. Numerous registration techniques have been proposed based on choosing a specific feature, deformation model, optimization method, and/or similarity measure. See [2] for a recent survey of image registration techniques.

For image registration, we need to recover the geometric transformation and/or intensity function. Let $I_1(u,v)$ and $I_2(x,y)$ be the reference and observed images, respectively. The relationship between these images is $I_1(u,v) = T_2\{I_2(T_1\{(x,y)\})\}$, where T_1 is a two-dimensional (2-D) geometric transformation operator that relates the (x,y) coordinates in I_2 to the (u,v) coordinates in I_1 and T_2 is the intensity function.

The estimation of the intensity function T_2 is useful when we want to register images taken from different sensors or when illumination is changed by automatic gain exposure of a camera. Comparametric equations have been introduced to model the intensity function T_2 [3]. Although these equations are nonlinear, a piecewise linear method has been developed to estimate T_1 and T_2 simultaneously [4].

Mutual information is a similarity measure that has recently been introduced for multimodal medical image registration [5], [6]. Correlation ratio is another similarity measure for multimodal image registration and has proven to perform better than mutual information [7]. Multimodal image registration has been studied extensively in the medical imaging domain. In this work, we assume that the intensity function is linear. Similarity measures like the zero-mean normalized sum of squared differences (SSD) and correlation coefficient are invariant to the linear intensity changes.

This paper describes a hierarchical image registration system. We model the mapping function as a perspective transformation. The algorithm estimates the perspective parameters necessary to register any two misaligned digital images. The parameters are selected to minimize the SSD between the two images. They are computed iteratively in a coarse-to-fine hierarchical framework using a variation of the Levenberg–Marquadt nonlinear least squares optimization method. This approach yields a robust solution that precisely registers images with subpixel accuracy.

The primary drawback of the optimization-based approach is that it may fail unless the two images are misaligned by a moderate difference in scale, rotation, and translation. In order to address this problem, we introduce a log-polar registration module to bring the images into approximate alignment, even in the presence of arbitrary rotation angles and large scale changes. Its purpose is to furnish a good initial estimate to the perspective registration module that is based on nonlinear least squares optimization.

The scope of this work shall prove useful for various applications, including the registration of aerial images, and the formation of image mosaics. Note that aerial imagery may be acquired from uncalibrated airborne cameras subjected to yaw, pitch, and roll at various altitudes. Since the terrain appears flat from moderately high altitude, it is an ideal candidate for registration using a single perspective transformation. An example demonstrating the registration of two aerial images in the presence of large scale/rotation and moderate perspective is shown in Fig. 1. The image in Fig. 1(a) is automatically registered to that in Fig. 1(b), as depicted by the highlighted rectangle.

In Section II, we discuss related work on the standard Levenberg–Marquardt algorithm (LMA) and log-polar techniques. Section III describes a modified LMA for improving the performance of the standard LMA and Section IV presents our proposed log-polar method. In Section V, we demonstrate the success of the log-polar transform in recovering large deformations by comparing registration accuracy with and without the log-polar registration module. A significant increase in correct matches is attributed to our algorithm. A secondary comparison was made by replacing the log-polar module with the well-known Fourier–Mellin transform. Again, our log-polar module proved superior to the Fourier–Mellin transform for achieving high perspective registration accuracy.

II. Previous Work

In this section, we discuss related work on the LMA and the log-polar techniques. In Section II-A, we present a background of the Levenberg–Marquardt nonlinear least-squares optimization algorithm that is useful for achieving subpixel registration accuracy. The log-polar transform is described in Section II-B.

where $\mathbf{H}(\mathbf{a})$ is the Hessian matrix and $\mathbf{B}(\mathbf{a})$ is the residual vector

$$b_k(\mathbf{a}) = \frac{\partial \chi^2(\mathbf{a})}{\partial a_k} = -2\sum_{i=1}^{N}\left[I_1(\mathbf{u}_i) - I_2'(\mathbf{u}_i)\right]\frac{\partial I_2'(\mathbf{u}_i)}{\partial a_k} \quad (6)$$

$$h_{kl} = \frac{\partial^2 \chi^2(\mathbf{a})}{\partial a_k \partial a_l} \approx 2\sum_{i=1}^{N}\left[\frac{\partial I_2'(\mathbf{u}_i)}{\partial a_k}\frac{\partial I_2'(\mathbf{u}_i)}{\partial a_l}\right]. \quad (7)$$

We can improve the standard Levenberg–Marquardt optimization algorithm outlined above by adding two modifications. The first modification includes the use of a multiresolution pyramid for both reference and target images. The second modification virtually eliminates the calculation of the Hessian matrix (7) which would otherwise have been computed in every iteration. Our second modification is based on the work of [16], whereby registration was performed on medical images subjected to similarity transforms (rotation/scale/translation). We have extended their method to recover perspective parameters.

A. Multiresolution Pyramid

A multiresolution pyramid consists of a set of images representing an image in multiple resolutions. The original image, sitting at the base of the pyramid, is downsampled by a constant scale factor in each dimension to form the next level. This is repeated from one level to the next until the tip of the pyramid is reached. The image size at level i is reduced from the original by a factor of 2^i in each dimension. Level 0, at the base of the pyramid, is referred to as the finest level. Level $n-1$, at the tip of the pyramid, is known as the coarsest level.

Multiresolution pyramids supply us with two major advantages. First, when we apply the Levenberg–Marquardt method to the coarsest level of the pyramid, the number of pixels is reduced by a factor of $2^{2(n-1)}$. We get large computational gains because most of the iterations are executed in the coarsest level, consisting of fewer pixels. Second, the smoothness conditions imposed by successively bandlimiting the pyramid levels causes $\chi^2(\mathbf{a})$ to be computed on smoother images. This smoothness property helps prevent getting trapped in local minimas. An example of $\chi^2(\mathbf{a})$ computed on two different pyramid levels is shown in Fig. 4. Since the coarsest level retains large-scale features only, the registration algorithm proceeds from the coarsest level to progressively finer levels, where small corrections due to finer details are integrated. This approach passes the computed parameters as an initial estimate to the next finer level. The parameters must be scaled properly across successive levels. Let the scale factor between the levels be s: $x^{i+1} = sx^i$, $y^{i+1} = sy^i$, $u^{i+1} = su^i$, and $v^{i+1} = sv^i$, where

$$u^i = \frac{a_1^i x^i + a_2^i y^i + a_3^i}{a_7^i x^i + a_8^i y^i + 1} \quad (8a)$$

$$v^i = \frac{a_4^i x^i + a_5^i y^i + a_6^i}{a_7^i x^i + a_8^i y^i + 1}. \quad (8b)$$

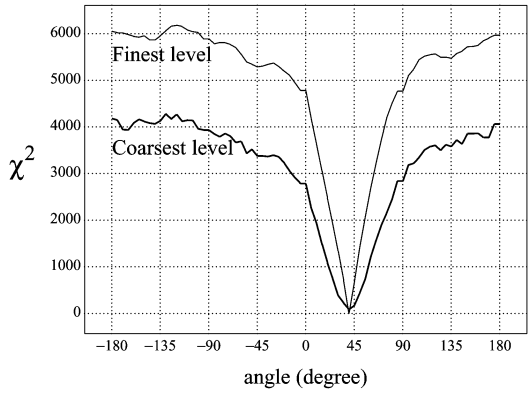

Fig. 4. Example of $\chi^2(\mathbf{a})$ computed on two different pyramid levels.

Substituting the coordinates of the next finer level into the above equations yields

$$\frac{u^{i+1}}{s} = \frac{a_1^i \frac{x^{i+1}}{s} + a_2^i \frac{y^{i+1}}{s} + a_3^i}{a_7^i \frac{x^{i+1}}{s} + a_8^i \frac{y^{i+1}}{s} + 1}$$
$$\frac{v^{i+1}}{s} = \frac{a_4^i \frac{x^{i+1}}{s} + a_5^i \frac{y^{i+1}}{s} + a_6^i}{a_7^i \frac{x^{i+1}}{s} + a_8^i \frac{y^{i+1}}{s} + 1}. \quad (9)$$

Multiplying both sides by s gives us

$$u^{i+1} = \frac{a_1^{i+1}x^{i+1} + a_2^{i+1}y^{i+1} + a_3^{i+1}}{a_7^{i+1}x^{i+1} + a_8^{i+1}y^{i+1} + 1} = \frac{a_1^i x^{i+1} + a_2^i y^{i+1} + sa_3^i}{a_7^i \frac{x^{i+1}}{s} + a_8^i \frac{y^{i+1}}{s} + 1}$$
$$v^{i+1} = \frac{a_4^{i+1}x^{i+1} + a_5^{i+1}y^{i+1} + a_6^{i+1}}{a_7^{i+1}x^{i+1} + a_8^{i+1}y^{i+1} + 1} = \frac{a_4^i x^{i+1} + a_5^i y^{i+1} + sa_6^i}{a_7^i \frac{x^{i+1}}{s} + a_8^i \frac{y^{i+1}}{s} + 1}. \quad (10)$$

Thus, the relation between parameters is

$$a_1^{i+1} = a_1^i \quad a_2^{i+1} = a_2^i \quad a_3^{i+1} = sa_3^i$$
$$a_4^{i+1} = a_4^i \quad a_5^{i+1} = a_5^i \quad a_6^{i+1} = sa_6^i$$
$$a_7^{i+1} = \frac{a_7^i}{s} \quad a_8^{i+1} = \frac{a_8^i}{s}. \quad (11)$$

In our case, $s = 2$, so the translation parameters a_3 and a_6 are multiplied by two and a_7 and a_8 divided by two.

B. Modified Levenberg–Marquardt Algorithm

In the standard LMA, we calculate the $\mathbf{B}_{8\times 1}$ vector and Hessian matrix $\mathbf{H}_{8\times 8}$ in each iteration. In this section, we review a modified LMA that realizes performance gains by eliminating the calculation of the Hessian matrix at each iteration. Consider the following objective function to establish a similarity measure between I_1 and I_2

$$\chi^2(\mathbf{a}) = ||I_1(\mathbf{u}) - I_2(T_{\mathbf{A}}\{\mathbf{x}\})||^2. \quad (12)$$

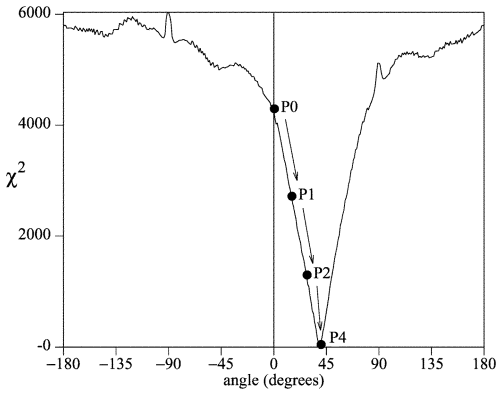

Fig. 5. χ^2 curve for rotation (standard LMA).

We shall assume that I_2 is mapped to I_1 after a series of perspective transformations $\mathbf{A}$. During the iterative process, new estimates for $\mathbf{A}$ are computed as follows:

$$\mathbf{A}_{i+1} = \mathbf{A}_i + \Delta\mathbf{A}_i \tag{13}$$

where $\mathbf{A}_i = \mathbf{I} + \Delta\mathbf{A}_1 + \Delta\mathbf{A}_2 + \cdots + \Delta\mathbf{A}_{i-1}$. Since I_2 is transformed in each iteration, the Hessian matrix must be recomputed because it is a function of the gradient of I_2. The Hessian matrix is responsible for computing the $\Delta\mathbf{A}$ terms above. Fig. 5 depicts the series of parameter estimates beginning from the initial guess P_0.

The goal of the modified LMA is to eliminate the computation of the Hessian matrix. This is achieved by casting this problem into one where I_1 is transformed into I_2, leaving I_2 unchanged from one iteration to the next. This permits the Hessian matrix to be computed only once, i.e., in the first iteration.

In order to determine the new estimates for $\mathbf{A}$ in the modified LMA, we must express χ^2 in terms of a transformation that maps I_1 to I_2. The fundamental difference between the standard LMA and the modified LMA is that the standard LMA updates the current estimate by changing the initial guess toward the global minima, while the modified LMA brings the global minima toward the initial guess. Fig. 6 depicts several snapshots of the χ^2 curve after 0, 10, 20, and 30 iterations, respectively. The consequence of this formulation can be summarized with the following update rule for the modified LMA:

$$\begin{aligned}\mathbf{A}_{i+1}^{-1} &= (\mathbf{I}+\Delta\mathbf{A}_i)^{-1}\cdots(\mathbf{I}+\Delta\mathbf{A}_2)^{-1}(\mathbf{I}+\Delta\mathbf{A}_1)^{-1}\\ &= (\mathbf{I}+\Delta\mathbf{A}_i)^{-1}\mathbf{A}_i^{-1}.\end{aligned} \tag{14}$$

An important distinction between the standard and modified LMA methods lie in the manner in which the unknown parameters are updated in each iteration. In the standard LMA, the initial estimates for the unknown parameters are chosen using identity matrix $\mathbf{I}$ as the initial guess for point P_0. Then, we calculate the directional derivatives of I_2, g_x and g_y, where $g_x = (\partial I_2/\partial x)$ and $g_y = (\partial I_2/\partial y)$. These processes are in the order of $O(N\times 3\times 3)$, where N is the number of pixels and 3×3 is the size of the kernel. The standard LMA gives us the $\Delta\mathbf{A}$ that we use to add to the initial guess $\mathbf{I}$ to move from point P_0 to P_1 on the $\chi^2(\mathbf{a})$ curve. In the next iteration, because the image I_2 is warped by $\mathbf{A}+\Delta\mathbf{A}$, we need to compute g_x and g_y again to find a new $\Delta\mathbf{A}$. Therefore, in the standard LMA, the optimal solution point P_i slides on the $\chi^2(\mathbf{a})$ curve. However, in the modified LMA, we shift the $\chi^2(\mathbf{a})$ curve toward the initial guess P_0. This is achieved by resampling I_1 with the inverse transformation $(\mathbf{I}+\Delta\mathbf{A}_i)^{-1}\mathbf{A}_i^{-1}$. Consequently, image I_1' is brought closer to I_2. The new image I_1' and image I_2 are now used to minimize $\chi^2(\mathbf{a})$. The result produces a new $\Delta\mathbf{A}$ that is always added to $\mathbf{I}$, the initial guess point P_0. Since I_2 does not change, we do not need to compute g_x and g_y. In Fig. 6, we see how the $\chi^2(\mathbf{a})$ graph is sliding toward the initial guess P_0.

We shall find it useful to rewrite χ^2 in terms of a forward mapping as well as an inverse mapping. This decomposition will enable us to apply a substantial part of the transformation to $I_1(\mathbf{u})$. As a result, the small inverse transformation that remains for $I_2(\mathbf{x})$ will permit us to drop the need to compute the Hessian. Suppose that we decompose the $T\{\}$ transformation into two transformations $T_{\mathbf{A}}T_{\mathbf{I}+\Delta\mathbf{A}}\{\}$. $T_{\mathbf{A}}\{\}$ is the transformation from the previous iteration and $T_{\mathbf{I}+\Delta\mathbf{A}}\{\}$ is the small transformation from the Levenberg–Marquardt method that minimizes $\chi^2(\mathbf{a})$

$$\chi^2(\mathbf{a}) = \|I_1(\mathbf{u}) - I_2(T_{\mathbf{A}}\{T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\}\})\|^2 \tag{15a}$$

$$= \frac{1}{|\mathbf{A}|}\|I_1(T_{\mathbf{A}^{-1}}\{\mathbf{u}\}) - I_2(T_{\mathbf{A}^{-1}}\{T_{\mathbf{A}}\{T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\}\}\})\|^2 \tag{15b}$$

$$= \frac{1}{|\mathbf{A}|}\|I_1(T_{\mathbf{A}^{-1}}\{\mathbf{u}\}) - I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\})\|^2 \tag{15c}$$

$$= \frac{1}{|\mathbf{A}(\mathbf{I}+\Delta\mathbf{A})|}\|I_1(T_{(\mathbf{I}+\Delta\mathbf{A})^{-1}\mathbf{A}^{-1}}\{\mathbf{u}\}) - I_2(\mathbf{x})\|^2$$

$$= \frac{1}{|\mathbf{A}(\mathbf{I}+\Delta\mathbf{A})|}\|I_1(T^{-1}\{\mathbf{u}\}) - I_2(\mathbf{x})\|^2. \tag{15d}$$

Equation (15) shows the necessary steps to transform $T_{\mathbf{A}}T_{\mathbf{I}+\Delta\mathbf{A}}\{\}$ from the I_2 coordinate system to the I_1 coordinate system with proper normalization. Instead of minimizing (15a), we minimize (15c) with respect to the parameters $\Delta\mathbf{A}$.

In the modified LMA, we need to derive $\partial I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\})/\partial\Delta\mathbf{a}$ and the update rule for each transformation parameter. We can decompose $\partial I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\})/\partial\Delta\mathbf{a}$ as follows:

$$\frac{\partial I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\})}{\partial\Delta\mathbf{a}} = \frac{\partial I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\})}{\partial\mathbf{x}}\frac{\partial\mathbf{x}}{\partial\Delta\mathbf{a}}. \tag{16}$$

Since transformation $\Delta\mathbf{a}$ is small, then $I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\}) \approx I_2$ and $(\partial\mathbf{x}/\partial\Delta\mathbf{a}) \approx (\partial\mathbf{u}/\partial\Delta\mathbf{a})$. This yields

$$\frac{\partial I_2(T_{\mathbf{I}+\Delta\mathbf{A}}\{\mathbf{x}\})}{\partial\Delta\mathbf{a}} \approx \frac{\partial I_2}{\partial\mathbf{x}}\frac{\partial\mathbf{x}}{\partial\Delta\mathbf{a}} \approx \frac{\partial I_2}{\partial\mathbf{x}}\frac{\partial\mathbf{u}}{\partial\Delta\mathbf{a}} \tag{17}$$

where $(\partial I_2/\partial\mathbf{x})$ is the gradient of I_2. Thus, $(\partial I_2(\mathbf{x})/\partial\Delta a_k)$ for the eight parameters are

$$\frac{\partial I_2}{\partial\Delta a_1} = \frac{\partial I_2}{\partial x}\frac{\partial u}{\partial\Delta a_1} = g_x\frac{x}{w}$$

$$\frac{\partial I_2}{\partial\Delta a_2} = \frac{\partial I_2}{\partial x}\frac{\partial u}{\partial\Delta a_2} = g_x\frac{y}{w}$$

$$\frac{\partial I_2}{\partial\Delta a_3} = \frac{\partial I_2}{\partial x}\frac{\partial u}{\partial\Delta a_3} = g_x$$

$$\frac{\partial I_2}{\partial\Delta a_4} = \frac{\partial I_2}{\partial y}\frac{\partial v}{\partial\Delta a_4} = g_y\frac{x}{w}$$

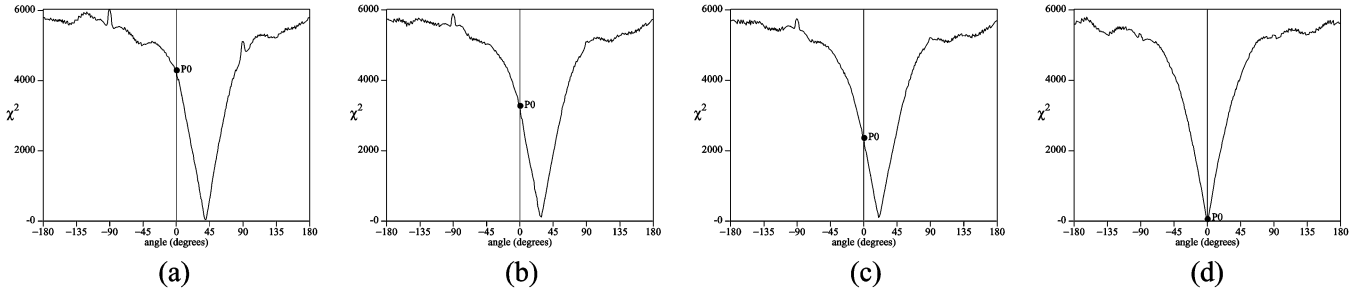

Fig. 6. χ^2 curve for rotation (modified LMA). (a) 0 iterations, (b) 10 iterations, (c) 20 iterations, and (d) 30 iterations.

$$\frac{\partial I_2}{\partial \Delta a_5} = \frac{\partial I_2}{\partial y}\frac{\partial v}{\partial \Delta a_5} = g_y \frac{y}{w}$$
$$\frac{\partial I_2}{\partial \Delta a_6} = \frac{\partial I_2}{\partial y}\frac{\partial v}{\partial \Delta a_6} = g_y$$
$$\frac{\partial I_2}{\partial \Delta a_7} = \frac{\partial I_2}{\partial x}\frac{\partial u}{\partial \Delta a_7} + \frac{\partial I_2}{\partial y}\frac{\partial v}{\partial \Delta a_7} = -x\left(g_x \frac{a_1x + a_2y + a_2}{w^2} + g_y \frac{a_4x + a_5y + a_6}{w^2}\right)$$
$$\frac{\partial I_2}{\partial \Delta a_8} = \frac{\partial I_2}{\partial x}\frac{\partial u}{\partial \Delta a_8} + \frac{\partial I_2}{\partial y}\frac{\partial v}{\partial \Delta a_8} = -y\left(g_x \frac{a_1x + a_2y + a_2}{w^2} + g_y \frac{a_4x + a_5y + a_6}{w^2}\right)$$

where $w = a_7x + a_8y + 1$. In the standard LMA, the update rule is as follows:

$$\mathbf{A}_{\text{new}} = \mathbf{A}_{\text{old}} + \Delta\mathbf{A}. \tag{18}$$

From (14), the updating rule for the modified LMA is

$$\mathbf{A}_{\text{new}} = \mathbf{A}_{\text{old}}(\mathbf{I} + \Delta\mathbf{A}) = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{bmatrix} \times \begin{bmatrix} \Delta a_1 + 1 & \Delta a_2 & \Delta a_3 \\ \Delta a_4 & \Delta a_5 + 1 & \Delta a_6 \\ \Delta a_7 & \Delta a_8 & 1 \end{bmatrix}. \tag{19}$$

In this implementation, we use a triangle filter for prefiltering the images to build the multiresolution pyramid. We used Fant's resampling algorithm to warp I_1 to I_2 using the estimated perspective transformation at each iteration. Note that Fant's algorithm uses linear interpolation (i.e., triangle filter) for reconstruction and unweighted averaging (i.e., box filter) for antialiasing. For further details about resampling, see [57]. The modified LMA version implemented in [16] uses a least-squared spline of order three to perform resampling. Compared to their modified LMA implementation, we realize identical parameter estimations with simpler resampling algorithms at twice the speed. Now, we present the pseudocode for our method.

Modified Levenberg–Marquardt Algorithm
Build multiresolution pyramid for images I_1 and I_2
Initialize parameters to the identity matrix
Initialize λ with a modest value, e.g., $\lambda = 0.01$
for $i = L$ to 0 **do** /*L is the coarsest pyramid level */
  Compute directional gradients: $\partial I_2^i/\partial x$ and $\partial I_2^i/\partial y$
  Compute the 8×8 Hessian matrix $\mathbf{H}$
  while ($|\Delta\chi^2(\mathbf{a})| > T_1$ or $\lambda < T_2$) **do**
    Apply transformation $T_{\mathbf{A}^{-1}}$ on I_1 in level i
    Compute vector $\mathbf{B}$
    Solve linear equations $\mathbf{H}\Delta\mathbf{a} = \mathbf{B}$ for $\Delta\mathbf{a}$
    Evaluate $\chi^2(\mathbf{a} + \Delta\mathbf{a})$
    if $\chi^2(\mathbf{a} + \Delta\mathbf{a}) < \chi^2(\mathbf{a})$ **then do**
      $\lambda \leftarrow \lambda/10$
      $\mathbf{a} \leftarrow \mathbf{a}(I + \Delta\mathbf{a})$
    else do
      $\lambda \leftarrow \lambda * 10$
    end if
  end while
end for

Although the parameter estimation method features subpixel accuracy, the two images to be registered must first be fairly close in scale (within a factor of two), rotation (within 45°), and translation. The purpose of the log-polar module, described in Section IV, is to account for large geometric transformations, bringing images into close alignment even in the presence of large (five-fold) scale changes, as well as arbitrary rotations and translations.

IV. Global Registration Using Log-Polar Transform

We have implemented a new algorithm for automatically finding the translation between both input images in the presence of large scale and rotation. We emphasize that our method does not compute the Fourier transform and does not use phase-correlation. The search space has four dimensions. The new method is based on multiresolution log-polar transformations to simultaneously find the best scale, rotation, and translation parameters. The coarse-to-fine multiresolution framework accelerates the process by permitting estimates computed in the low resolution images to serve as initial guesses to the higher resolution images. We limit the search space to a small neighborhood about the initial guess. The size of that neighborhood shrinks as we move from the coarsest level to the finest level of the pyramid (e.g., from 8×8 to 2×2 search space). One of the benefits of the discrete log-polar transform is that we quantize the scale and rotation axes. Therefore, we have a finite number of points to search and this number is small at the coarsest level.

We crop a circular template from the reference image and compute its log-polar transformation. The radius and the center of the template are optionally given by the user. The radius

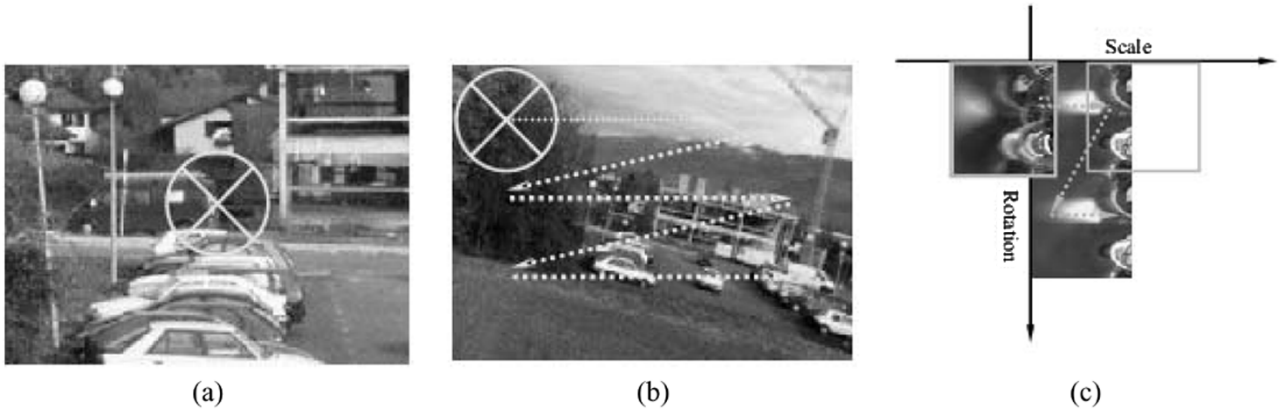

Fig. 7. Four-dimensional search strategy. (a) A circular template from the center of the reference image is cropped. (b) For every position in the target image, a circular region is selected and compared against the circular template in (a) to find the best (T_x, T_y). (c) Search for (R, S) in the log-polar domain.

varies from 25% to 10% of the image width Fig. 7(a). The default value for the radius is 25% of the input image width and the center of the template is the center of the reference image. Then, for all positions in the target image, we crop a circular image and compute the log-polar transformation Fig. 7(b). In the log-polar space, we map $[0° \cdots 360°]$ to $[0 \cdots h]$, where h is the height of the template. Note that we pad the template by wrapping around Fig. 7(c). We compute the *base* of the logarithm for log-polar transformation as follows:

$$\text{base} = e^{(\log w/w)} \tag{20}$$

where w is the width of the input image ($w = h =$ diameter). The choice of the base is arbitrary. However, we compute the base in (20) to set the width of the log-polar image to that of the input image. We can compute the maximum translation in log-polar space as follows: $(\log S_{\max}/\log \text{base})$. We set the maximum scale factor $S_{\max}$ to 5.0 to limit the search in the scale direction.

We have extensively tested several similarity measures, including normalized correlation coefficient, phase correlation, and mutual information. We use the normalized correlation coefficient similarity measure (21) due to its superior performance. It should be noted that mutual information may suffice for multimodal registration of MRI and PET scans in medical applications, whereby only small deformations are found. Our domain consists of images subjected to large similarity transformations acquired in one modality. The normalized correlation coefficient similarity measure is given as follows:

$$\langle I_1 \circ I_2 \rangle = \frac{\sum (I_1 - \mu_1)(I_2 - \mu_2)}{\sqrt{\sum (I_1 - \mu_1)^2 \sum (I_2 - \mu_2)^2}} \tag{21}$$

where μ is the average of image I_i. The approach at any given level is outlined as follows.

1) Crop central region I_1' from I_1.
2) Compute I_{1p}', the log-polar transformation of I_1'.
3) For all positions (x, y) in I_2:
4) Crop region I_2'.
5) Compute I_{2p}'.
6) Cross correlate I_{1p}' and $I_{2p}' \rightarrow (dx, dy)$.
7) If maximum correlation, save (x, y) and (dx, dy).
8) Scale $\leftarrow dx$.
9) Rotation $\leftarrow dy$.
10) Translation $\leftarrow (x, y)$.

The procedure outlined above recovers the origin of the log-polar transform, as well as the global scale and rotation.

In our implementation, a pair of 640×480 images are registered in approximately 20 seconds on a 3.06-GHz Pentium 4 machine. The computational complexity of our algorithm is $O(N^2)$, where N is the resolution of the coarsest level of the pyramid. The user typically sets this resolution to be 64×64 or 32×32. Although techniques such as the Fourier–Mellin transform exploit the $O(N \log N)$ complexity of the FFT to efficiently find the log-polar origin, our method is local and is thereby more robust to projective transformations and large scale changes. Furthermore, the bulk of our computation is performed at the coarsest level where there are fewest pixels.

V. Experimental Results

An analytical evaluation of the robustness of image registration algorithms is an elusive task. Performance is highly dependent on the content of the input images. Although image models may exist for particular domains, the deformations, and noise functions that may apply to images defy restrictive bounds. Consequently, many proposed image registration algorithms in the literature have limited their published results to the use of a few reference images and their synthetically generated target images. In an effort to broaden our test suite, we chose an empirical approach with a variety of input images. In Section V-A, we demonstrate the registration of images subjected to large changes in scale and rotation. The reference and target images are taken by a camera with optical zoom. In Section V-B, we test the robustness of our algorithm with a large suite of 10 000 image pairs.

A. Uncalibrated Test Images

A Canon PowerShot G3 digital camera with 4× optical zoom was used to capture a set of test images taken from natural and man-made scenes. We took 30 pairs of uncalibrated images without a tripod. The content of these images varies from very highly textured to minimally textured areas. Several of our test images are problematic for feature-based methods, since these images have smooth surfaces with no distinctive features. An

Fig. 8. (a) Observed images: 4× zoom, arbitrary rotation, and moderate perspective. (b) Registration results highlighted on reference images.

example of such a pair is shown in Fig. 1. Our method uses all pixels and does not depend on any specific feature set. Images were acquired with (a) no magnification and (b) 4× magnification with unknown rotation about the optical axis. We have identified the boundaries of the magnified images in their counterpart target images with an overlaid white rectangle. We do not have ground truth transformation parameters for the image pair. In order to quantify registration accuracy, we compute the correlation coefficient in the overlapping area. The correlation coefficient values for the thirty pairs mentioned above are all above 0.9, which is very good considering camera noise and artifacts introduced by warping to produce the target images. The resulting overlays demonstrate high accuracy, as depicted by the absence of any noticeable visual misalignment. Five pairs of images from the set are shown in Fig. 8.

We had access to source code for several registration methods, including those of Lowe (SIFT) [49], Georgescu/Meer (CDOF) [52], and Mann/Picard (Video Orbits) [18], and Li/Manjunath/Mitra (contour-based) [42]. All of the non-SIFT methods failed to register the image pairs in Fig. 8. They generally failed in all of our examples involving large optical zoom. Only the SIFT method produced competitive results. In Fig. 8, for instance, the SIFT method misregistered the roof and car examples, while the other methods misregistered all the images in the figure.

B. Calibrated Test Images

It is not feasible to capture a very large set of images with a variety of image content and transformation parameters. Therefore, we tested the robustness of our algorithm with 10 000 image pairs whose transformations are known. We sampled 1000 images from the Corel Stock Photo Library for our test images. The Corel library contains 20 000 royalty-free photographic images on 200 CD-ROM's (http://www.corel.com/). Each CD contains a different category of images. We randomly sampled five images from each CD, for a total of 1,000 images. Then, we randomly generated ten different sets of perspective parameters for each image. In this manner, we uniformly sampled 10 000 points from the parameter space. The range of these parameters are as follows: $\alpha \in [-30^\circ \cdots 30^\circ]$, $\beta \in [-30^\circ \cdots 30^\circ]$, $\gamma \in [0^\circ \cdots 180^\circ]$, $s \in [1.0 \cdots 4.5]$, $t_x \in [-40 \cdots 40]$, $t_y \in [-40 \cdots 40]$, where γ is the rotation about the optical axis, s is the scale factor of the digital zoom, and α and β rotate the image plane about the x—and y—axes, respectively. These rotations introduce foreshortening effects. We generated 10 000 target images from these random parameters. We set up four experiments as follows.

1) Log-Polar/LMA (LP/LMA).
2) Fourier–Mellin/LMA (FM/LMA).
3) Only LMA (LMA).
4) SIFT.

First, we used our log-polar module to recover the global rotation, scale, and translation parameters. The estimated parameters serve as the initial guess for the LMA module, which finds the perspective parameters with subpixel accuracy. The registration operations conducted over the set of 10 000 image pairs yielded a 93.70% success rate. In order to quantify the accuracy of the eight estimated perspective parameters, we calculated correlation coefficient values between the actual parameters $\mathbf{a}_{\text{act}}$ and the estimated parameters $\mathbf{a}_{\text{est}}$. This is akin to computing the cosine of the angle between the two eight-element vectors associated with the source and target images

$$\rho = \frac{\mathbf{a}_{\text{act}}\mathbf{a}_{\text{est}}^T}{\|\mathbf{a}_{\text{act}}\|\|\mathbf{a}_{\text{est}}\|}. \tag{22}$$

These values are between $[-1 \cdots 1]$. If the correlation value is close to one, it means the error between actual and estimated parameters is very small. The histogram of the errors is shown in Fig. 10, where a log-scale is used along the vertical axis. The histogram depicts a sharp peak above 0.8 and the majority (93.70%) of the cases are concentrated in this section. Therefore, our measure of success for alignment is a correlation coefficient that exceeds 0.8. The main rationale for the 6.30% mismatch is related to the content of several images in the test set. In particular, these images had large uniform intensity

Fig. 9. Registration failed when applied to database images that do not have discriminating information in the central region, as shown above.

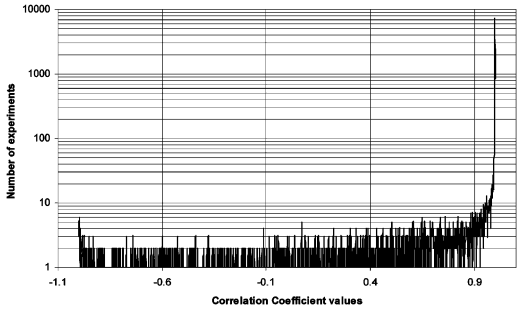

Fig. 10. Histogram of ρ for the LP/LMA case.

(smooth areas) in the central region of the image from which the log-polar algorithm crops a template window. Lack of visual information there leads the algorithm to false matches. We have shown two of these problematic images in Fig. 9.

In the second experiment, we use the Fourier–Mellin method to recover the global rotation, scale, and translation parameters that serve as initial estimates to the LMA module. In this case, we obtained a 71.24% success rate. Notice that our log-polar module yields fewer mismatches than the Fourier–Mellin transform. The main reason for these results is the perspective effects introduced by rotating the image with angles α and β. The Fourier–Mellin transform is a global method. Nonlinear sampling alters the Fourier coefficients, rendering the linear properties of the Fourier transform invalid. Our method uses a local circular region from the reference image and searches for its counterpart region in the target image, as shown in Fig. 7. This local approach permits our method to be less sensitive to perspective effects because perspective can be approximated to be locally affine [20].

In order to show the importance of the log-polar module, we ran the LMA without the estimated initial parameters from the log-polar module. Not surprisingly, the LMA module performs poorly when no initial estimate is provided. Only a 28.77% success rate was achieved. The histogram of the error is plotted in Fig. 11, whereby the vertical axis is represented in a log-scale.

Since the SIFT method [49] produced competitive results in our uncalibrated dataset, we applied that leading registration technique to our 10 000 image pairs. Note that our implementation of SIFT came directly from the source code of Lowe and Brown. SIFT performed well, yielding an alignment success rate of 87.97%. Our method remains superior at no additional computational cost. We summarize the results of the four experiments in Table I.

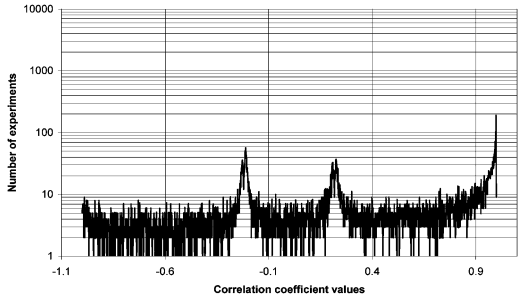

Fig. 11. Histogram of ρ for the LMA case.

TABLE I
REGISTRATION RESULTS FOR 10 000 IMAGE PAIRS

Method	Success rate
LP/LMA	93.70%
FM/LMA	71.24%
LMA	28.77%
SIFT	89.97%

C. Image Mosaics

We have tested our registration algorithm to create image mosaics by stitching together low resolution frames from several overlapping images. Mosaic techniques have been used to render images with large fields of view. The process involves two steps: 1) alignment of frames in the sequence and 2) composition (blending) of these aligned frames. Remarkable progress has been documented during the last decade in this area [12], [14], [18], [19], [58], [59]. In Fig. 12, a set of input images is presented. The set was acquired from http://www.inrialpes.fr/movi/. Notice that the images differ by large transformations in perspective, scale, and rotation. Our algorithm produced the image mosaic shown in Fig. 13. In order to best expose any misalignment, we applied unweighted averaging upon the overlapping areas. No advanced feathering technique was used since they can be misleading through their ability to hide minor misalignments.

For those applications that require seamless mosaics, we implemented a blending function in which the pixels in the overlapping area are weighted based on their distance from their respective image centers. Blending techniques vary from the classical work described in [60] to the more recent work in [61]. The weighting function we selected consists of a cosine fall-off that more heavily weighs the central pixels (23)

$$\begin{aligned} d_1 &= \sqrt{(x_1 - x_c)^2 + (y_1 - y_c)^2} \\ d_2 &= \sqrt{(x_2 - x_c)^2 + (y_2 - y_c)^2} \\ w_1 &= f(d_1), \quad w_2 = f(d_2) \\ I_{out} &= \frac{w_1 1 * I_1 + w_2 * I_2}{w_1 + w_2} \end{aligned}$$

where $f(t)$ is

$$f(t) = \frac{1 + \cos(t)}{2}. \tag{23}$$

Fig. 12. Image set used to create a panorama image.

Fig. 13. Mosaic produced by stitching images shown in Fig. 12.

VI. Conclusion and Future Work

This paper has presented a hierarchical image registration algorithm to register any two digital images misaligned due to large similarity transformation and mild perspective deformation. We introduce a new approach based on log-polar transforms in the spatial domain to handle the similarity transformation consisting of rotation, scale, and translation. Next, the algorithm couples the log-polar transform with a nonlinear least squares algorithm to estimate the perspective transformation parameters. Although the Fourier–Mellin transform also uses the log-polar transformation to recover rotation and scale, it is applied to the frequency domain only and is limited in use to small scale factors (< 2.0). Larger scale changes induce too much distortion to the Fourier coefficients to be useful for recovering the rotation, scale, and translation parameters. Instead, our work operates directly in the spatial domain and simultaneously recovers the best rotation, scale, and translation by performing correlation on tiles that have been transformed into log-polar space.

The purpose of the log-polar registration module is to bring two images into alignment using only rotation, scale, and translation. This serves as a fine estimate for the subsequent perspective registration module based on nonlinear least squares optimization. That module, based on the LMA, offers subpixel precision. Coupling the two modules in this manner facilitates the registration of images in the presence of large-scale $(\leq 5\times)$ and moderate perspective transformations.

Image resampling poses the largest bottleneck for the log-polar or Levenberg–Marquardt optimization modules. Currently, our image resampling function is implemented in software. With the advent of fast graphic cards and specialized GPU's, we will investigate the use of hardware resampling for improving the performance of our algorithms.

Additional future work will accelerate correlation. Currently, correlation in the log-polar domain consists of sliding a cropped window in raster order. It is worthwhile to examine whether this process may be accelerated by positioning the sliding window on areas of high information content only. Entropy, variance, or other statistically discriminating techniques can be used to quantify information content. Recent success with scale-invariant interest points (e.g., SIFT) suggest that the log-polar windows should be centered at these extracted positions. The reduction of correlation sites in the spatial domain should serve to eliminate unncessary log-polar warping and subsequent correlation in the log-polar domain.

Acknowledgment

The authors would like to thank Prof. R. Haralick and Dr. C. Weiman for their perceptive comments. They would also like to thank G. Yu for providing experimental data sets.

References

[1] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1992.
[2] B. Zitova and J. Flusser, "Image registration methods: A survey," *IVC*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.
[3] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1389–1406, Aug. 2000.
[4] F. M. Candocia, "Jointly registering images in domain and range by piecewise linear comparametric analysis," *IEEE Trans. Image Process.*, vol. 12, no. 4, pp. 409–419, Apr. 2003.
[5] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1995.
[6] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Seutens, and G. Marchal, "Automated multimodality image registration using information theory," in *Proc. 14th Int. Conf. Information Processing in Medical Imaging*, 1995, pp. 263–274.
[7] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Proc. 1st Int. Conf. Medical Image Computing and Computer-Assisted Intervention*, Oct. 1998, pp. 1115–1124.
[8] J. R. Bergen and E. Adelson, "Hierarchical, computationally efficient motion estimation algorithm," *J. Opt. Soc. Amer.*, vol. 4, no. 35, 1987.
[9] P. Anandan, "A computational framework and an algorithm for the measurement of visual motion," *Int. J. Comput. Vis.*, vol. 2, pp. 283–310, 1989.
[10] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proc. Eur. Conf. Computer Vision*, 1992, pp. 237–252.
[11] S. E. Chen, "Quicktime VR: An image-based approach to virtual environment navigation," in *Proc. Computer Graphics*, 1995, pp. 29–38.
[12] M. Irani and P. Anandan, "Video indexing based on mosaic representations," *Proc. IEEE*, vol. 86, no. 5, pp. 237–252, May 1998.
[13] R. Szeliski, "Image mosaicing for tele-reality applications," in *Proc. IEEE Workshop on Applications of Computer Vision*, 1994, pp. 230–236.
[14] R. Szeliski and H.-Y. Shum, "Video mosaics for virtual environments," in *Proc. IEEE Computer Graphics and Applications*, vol. 16, 1996, pp. 22–30.

[15] ——, "Creating full view panoramic image mosaics and environment maps," in *Proc. Computer Graphics*, 1997, pp. 251–258.
[16] P. Thévenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 27–41, Jan. 1998.
[17] H.-Y. Shum and R. Szeliski, "Construction and refinement of panoramic mosaics with global and local alignment," in *Proc. Int. Conf. Computer Vision*, 1998, pp. 953–958.
[18] S. Mann and R. W. Picard, "Video orbits of the projective group: A simple approach to featureless estimation of parameters," *IEEE Trans. Image Process.*, vol. 6, no. 9, pp. 1281–1295, Sep. 1997.
[19] J. Davis, "Mosaics of scenes with moving objects," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1998, pp. 354–360.
[20] G. Wolberg and S. Zokai, "Image registration for perspective deformation recovery," in *Proc. SPIE Conf. Automatic Target Recognition X*, vol. 4050, Apr. 2000.
[21] W. H. Marshall, C. N. Woolsey, and P. Bard, "Observations on cortical somatic sensory mechanisms of cat and monkey," *J. Neurophysiol.*, vol. 4, pp. 1–43, 1941.
[22] D. H. Hubel and T. N. Wiesel, "Sequence regularity and geometry of orientation columns in the monkey striate cortex," *J. Comp. Neurol.*, pp. 158–267 293, 1974.
[23] E. L. Schwartz, "Topographical mapping in primate visual cortex: History, anatomy and computation," in *Visual Science and Engineering. Models and Applications*, D. K. editor, Ed. New York: Marcel-Dekker, 1994.
[24] ——, "Spatial mapping in primate sensory projection: Analytic structure and relevance to perception," *Biol. Cybern.*, vol. 25, pp. 181–194, 1979.
[25] D. Weinshall and E. L. Schwartz, "A new method for measuring the visuotopic map function of striate cortex: Validation with macaque data and possible extension to measurement of the human map," *Soc. Neurosci. Abstr.*, pp. 1291–1291, 1987.
[26] S. W. Wilson, "On the retino-cortical mapping," *Int. J. Man-Mach. Studies*, vol. 18, pp. 361–389, 1983.
[27] F. Ferrari, J. Nielsen, and G. Sandini, "Space variant imaging," *Sens. Rev.*, vol. 15, no. 2, pp. 17–20, 1995.
[28] R. Wodnicki, G. W. Roberts, and M. Levine, "Design and evaluation of a log-polar image sensor fabricated using a standard 1.2 μm asic cmos process," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1274–1277, 1997.
[29] P. C. Arribas and F. M. Maciá, "Fpga implementation of a log-polar algorithm for real time applications," in *Proc. Conf. Design of Circuits and Integrated Systems*, 1997, pp. 63–68.
[30] F. J. Blasco, F. Pardo, and J. A. Boluda, "A fpga based pci bus interface for a realtime log-polar image processing system," in *Proc. Conf. Design of Circuits and Integrated Systems*, Nov. 1999, pp. 379–384.
[31] R. Etienne-Cummings, J. V. der Spiegel, P. Mueller, and M. Z. Zhang, "A foveated silicon retina for two-dimensional tracking," *IEEE Trans. Circuits Syst. II, Analog. Digit. Signal Process.*, vol. 47, pp. 504–517, Jun. 2000.
[32] D. Casasent and D. Psaltis, "Position, rotation, and scale-invariant optical correlation," *Appl. Opt.*, vol. 15, pp. 1793–1799, 1976.
[33] E. D. Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 3, pp. 700–703, Sep. 1987.
[34] Q. S. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier–Mellin transforms for image registration and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 12, pp. 1156–1168, Dec. 1994.
[35] B. S. Reddy and B. N. Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 5, no. 8, pp. 1266–1270, Aug. 1996.
[36] L. Lucchese, G. Cortelazzo, and M. Rizzato, "A phase correlation technique for estimating planar rotations," in *Proc. Int. Workshop Time-Varying Image Processing and Moving Object Recognition*, Sep. 1996, pp. 244–249.
[37] S. H. Chang, F. H. Chen, W. H. Hsu, and G. Z. Wu, "Fast algorithm for pointer pattern matching; invariant to translations, rotations, and scale changes," *Pattern Recognit.*, vol. 30, no. 2, pp. 311–320, Feb. 1997.
[38] L. Lucchese and G. M. Cortelazzo, "Noise-robust estimation of planar roto-translations with high precision," in *Proc. IEEE Int. Conf. Image Processing*, Oct. 1997, pp. 699–702.
[39] L. Lucchese, G. M. Cortelazzo, and C. Monti, "High resolution estimation of planar rotations based on fourier transform and radial projections," in *Proc. ISCAS*, vol. 2, Jun. 1997, pp. 1181–1184.
[40] H. Stone, M. Orchard, and E. Chang, "Subpixel registration of images," in *Proc. 33rd Asilomar Conf. on Signals, Systems, and Computers*, Oct. 1999, pp. 1446–1452.
[41] H. Stone, B. Tao, and M. McGuire, "Analysis of image registration noise due to rotationally dependent aliasing," *J. Vis. Commun. Image Represen.*, vol. 14, pp. 114–135, 2003.
[42] H. Li, B. S. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 320–334, Mar. 1995.
[43] J. Flusser and T. Suk, "A moment-based approach to registration of images with affine geometric distortion," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 2, pp. 382–387, Mar. 1994.
[44] X. Dai and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2351–2362, Sep. 1999.
[45] A. D. Ventura, A. Rampini, and R. Schettini, "Image registration by recognition of corresponding structures," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 3, pp. 305–314, May 1990.
[46] M. Holm, "Toward automatic rectification of satellite images using feature based matching," in *Proc. Int. Geoscience and Remote Sensing Symp.*, 1991, pp. 2439–2442.
[47] A. Goshtasby, G. C. Stockman, and C. V. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geosci. Remote Sens.*, vol. 24, no. 3, pp. 390–399, May 1986.
[48] M. Brown and D. G. Lowe, "Recognizing panoramas," in *Proc. Int. Conf. Computer Vision*, Oct. 2003, pp. 1218–1225.
[49] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[50] F. Schaffalitzky and A. Zisserman, "Viewpoint invariant texture matching and wide baseline stereo," in *Proc. ICCV*, Jul. 2001, pp. 636–643.
[51] T. Tuytelaars and L. V. Gool, "Wide baseline stereo matching based on local affinely invariant regions," in *Proc. 11th British Machine Vision Conf.*, 2000, pp. 412–422.
[52] B. Georgescu and P. Meer, "Point matching under large image deformations and illumination changes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 6, pp. 674–689, Jun. 2004.
[53] V. Ferrari, T. Tuytelaars, and L. V. Gool, "Wide-baseline multiple-view correspondences," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2003, pp. 718–725.
[54] Y. Dufournaud, C. Schmid, and R. Horaud, "Matching images with different resolutions," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Jun. 2000, pp. 612–618.
[55] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Proc. Int. Conf. Computer Vision*, Jul. 2001, pp. 525–531.
[56] Y. Dufournaud, C. Schmid, and R. Horaud, "Image matching with scale adjustment," INRIA Rhone-Alpes Montbonnot Saint-Martin, Tech. Rep. RR 4458, 2002.
[57] G. Wolberg, *Digital Image Warping*. Los Alamitos, CA: IEEE Computer Soc., 1990.
[58] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models Image Process.*, vol. 53, pp. 231–239, Mar. 1991.
[59] S. Peleg and J. Herman, "Panoramic mosaicing with videobrush," in *Proc. DARPA Image Understanding Workshop*, May 1997, pp. 261–264.
[60] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. Graphics*, vol. 2, no. 4, pp. 217–236, Oct. 1983.
[61] L. Wang, S. Kang, R. Szeliski, and H. Shum, "Optimal texture map reconstruction from multiple views," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Dec. 2001, pp. 347–354.

Siavash Zokai received the B.S. degree in electrical engineering from the Sharif University of Technology in 1990 and the M.S. and Ph.D. degrees in computer science from the City College of New York in 1997 and 2004, respectively.

During 2002 and 2003, he was an intern and Consultant in the Imaging and Visualization Department, Siemens Corporate Research, Princeton, NJ. He is currently a Research Scientist at Brainstorm Technology LLC, New York. His research interests include image registration, 3-D photography, and augmented reality.

George Wolberg (S'83–M'90–SM'98) received the B.S. and M.S. degrees in electrical engineering from Cooper Union, New York, in 1985 and the Ph.D. degree in computer science from Columbia University, New York, in 1990.

He is a Professor of computer science at the City College of New York. His research interests include image processing, computer graphics, and computer vision. He is the author of *Digital Image Warping* (Los Alamitos, CA: IEEE Computer Society Press, 1990), the first comprehensive monograph on image warping and morphing.

Prof. Wolberg is the recipient of a 1991 NSF Presidential Young Investigator Award, the 1997 City College of New York Outstanding Teaching Award, and the 2000 NYC Mayor's Award for Excellence in Science and Technology.