

Local Image Registration by Adaptive Filtering

Gulcin Caner, *Student Member, IEEE*, A. Murat Tekalp, *Fellow, IEEE*, Gaurav Sharma, *Senior Member, IEEE*, and Wendi Heinzelman, *Senior Member, IEEE*

Abstract—We propose a new adaptive filtering framework for local image registration, which compensates for the effect of local distortions/displacements without explicitly estimating a distortion/displacement field. To this effect, we formulate local image registration as a two-dimensional (2-D) system identification problem with spatially varying system parameters. We utilize a 2-D adaptive filtering framework to identify the locally varying system parameters, where a new block adaptive filtering scheme is introduced. We discuss the conditions under which the adaptive filter coefficients conform to a local displacement vector at each pixel. Experimental results demonstrate that the proposed 2-D adaptive filtering framework is very successful in modeling and compensation of both local distortions, such as Stirmark attacks, and local motion, such as in the presence of a parallax field. In particular, we show that the proposed method can provide image registration to: a) enable reliable detection of watermarks following a Stirmark attack in nonblind detection scenarios, b) compensate for lens distortions, and c) align multiview images with nonparametric local motion.

Index Terms—Adaptive filtering, image registration, local image registration, nonparametric image registration, stirmark recovery, watermark synchronization.

I. INTRODUCTION

IMAGE registration plays a critically important role as a pre-processing step in many image processing and computer vision applications. A large number of techniques have been developed to solve different variants of this problem. Zitova *et al.* [1] present a survey of recent image registration techniques covering different application areas. One of the important factors to achieving accurately registered images is the model that describes the (spatial) mapping between the images to be registered. We can classify image registration techniques into two main groups: 1) global image registration methods that employ parametric spatial transformations and 2) local image registration methods that can handle spatially varying deformations. Global parametric transformations include translation, affine, pseudoperspective and perspective (homography) mappings [2],

[3]. These global transformations are valid under restrictive assumptions on the images being registered. For example, in multiview images the popular 8-parameter perspective mapping can accurately model only planar three-dimensional (3-D) scenes in the presence of arbitrary camera motion or arbitrary static scenes in the presence of only rotational camera motion (no translation) [4]. Parameters of these transformations can be estimated using either image gradients (direct methods) [5], [6] or feature-based methods [7], [8]. The accuracy of global registration methods proves insufficient when the underlying assumptions are not valid; hence, the globally modeled parametric motion shows systematic local deviations. In these scenarios, an alternative local registration method is required.

Local image registration methods are motivated by two main applications: 1) correction of locally varying image distortions, such as random bending attacks, e.g., Stirmark [9] and spatially-varying lens distortions, and 2) compensation of locally varying motions in the presence of a parallax field. Local image registration methods include dense motion estimation (optical flow) approaches [10]–[12], two-dimensional (2-D) mesh-based approaches [13], and methods based on an underlying 3-D scene representation [14]. Dense motion estimation methods make no assumption about the scene geometry and camera model, and rely only on image gradients for local registration using a motion (optical flow) field. 2-D mesh based approaches segment the image into triangular or quadratic patches, where a single parametric model is assumed to be valid for each patch [15]. Methods employing a 3-D scene representation include the plane + parallax framework [16]. These methods tend to be computationally complex and sensitive to errors in motion field estimation. Local image registration methods have also been employed to compensate for geometric distortion attacks, such as those intended to disable watermarks. These include transform-domain approaches [17], feature-based methods [18], and direct techniques [19]. In the context of image restoration, recently Šroubek *et al.* [20] have also proposed integration of registration into the restoration process for translational misregistration.

In this paper, we present a new local image registration technique based on adaptive filtering, for both space-varying distortion and motion compensation, which does not require explicit estimation of the local distortion/displacement field. Adaptive filters have been successfully applied to a number of one-dimensional (1-D) system-identification problems, such as echo-cancellation [21]. In these applications, adaptive filters not only allow for the estimation of an unknown system but also incorporate the capability to track smoothly varying changes in the system. In this paper, we formulate local image registration as a 2-D system identification problem with spatially varying system parameters. Since the successive update procedure in

Manuscript received August 4, 2005; revised February 8, 2006. This work was supported in part by the National Science Foundation ECS Sensors program under Grant ECS-0428157. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Thierry Blu.

G. Caner and W. Heinzelman are with the Electrical and Computer Engineering Department, University of Rochester, Rochester, NY 14627-0126 USA (e-mail: caner@ece.rochester.edu; gsharma@ece.rochester.edu; wheinzel@ece.rochester.edu).

A. M. Tekalp is with the College of Engineering, Koc University, Istanbul, Turkey (e-mail: mtekalp@ku.edu.tr).

G. Sharma is with the Electrical and Computer Engineering Department and the Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY 14627-0126 USA (e-mail: gsharma@ece.rochester.edu).

Digital Object Identifier 10.1109/TIP.2006.877514

the adaptive filtering is inherently 1-D, we map the 2-D image plane into a 1-D sequence using space-filling curves. This ensures spatial contiguity in the 2-D image plane, which is a prerequisite for filter convergence and tracking. The proposed method is computationally simpler than other approaches for local image registration. Image distortion/motion compensation using an adaptive filtering framework is presented in Section II. Section III discusses the conditions under which adaptive filter conforms to an underlying local displacement field. Application scenarios are discussed in Section IV together with extensive experimental results. The application scenarios include: 1) correction of Stirmark random bending attacks [22], 2) correction of camera/lens distortions, 3) registration of multiple views of a scene with local motion that is not representable as a parametric model. Conclusions are presented in Section V.

II. DISTORTION/MOTION COMPENSATION BY ADAPTIVE FILTERING

A. Local Distortion Modeling

Consider a pair of images $I_1(x, y)$ and $I_2(x, y)$ with overlapping views of the same scene, but with the underlying geometry, i.e., (x, y) spatial coordinates, of one image locally distorted or warped with respect to the other. Such images can be encountered in a number of applications; for instance, a watermarked image and a warped version obtained by applying random geometric distortions in order to attack the watermark (e.g., using Stirmark [9]). Other examples include scenes with peripheral overlap captured for the purpose of mosaicking where camera lens distortions induce different distortions in the geometry of overlapped regions [5], multiple views of a 3-D scene under (small) camera displacements, and a reproduction of an image captured through a print and scan process [23].

In image registration problems, the relation between the two images, over the region of overlap, is traditionally represented by

$$I_2(x, y) = I_1(x + d_x(x, y), y + d_y(x, y)) \quad (1)$$

where $\mathbf{d}(x, y) = (d_x(x, y), d_y(x, y))$ is a subpixel-valued spatially varying displacement field. The image registration problem can then be regarded as estimation of this spatially varying subpixel displacement field and mapping of image $I_2(x, y)$ onto the coordinate system of image $I_1(x, y)$.

In this work, we present an alternative model to represent local image distortions, where the mapping between $I_1(x, y)$ and $I_2(x, y)$ is expressed by means of a spatially varying linear filter $h_o(x, y; x_o, y_o)$, over the region of overlap as

$$I_2(x_o, y_o) = \sum_{x, y} h_o(x, y; x_o, y_o) I_1(x, y) + e(x_o, y_o). \quad (2)$$

The problem of image registration can now be regarded as a system identification problem, where the system response $h_o(\cdot)$ needs to be determined in order to map image $I_2(x, y)$ onto the coordinate system of image $I_1(x, y)$. In

the following, we propose an adaptive filtering framework for the estimation of $h_o(x, y; x_o, y_o)$ without explicitly estimating the subpixel-valued local displacement field $\mathbf{d}(x, y) = (d_x(x, y), d_y(x, y))$.

B. Local Image Registration as a 2-D System Identification Problem

Adaptive filters have been used extensively for system identification for 1-D (temporal) systems [21, Ch. 14]. Thus, they are good candidates for use with our proposed model of (2). In order to establish the context and to highlight the differences in our subsequent development, first consider adaptive filtering for 1-D temporal signals, which typically consists of a two step process: 1) a filtering step, where the filter coefficients, $\hat{h}(t, t_o)$ are convolved with the input signal, $v_1(t)$, to produce an estimate of the desired response, $v_2(t_o)$, and 2) an adaptive process where the set of filter coefficients are adjusted using the resulting estimation error, $e(t_o)$. For the commonly used least-mean-square (LMS) [21, Ch. 5] adaptation algorithm, the adaptive filtering process is given as

$$\hat{v}_2(t_o) = \sum_{(t \in U)} \hat{h}(t, t_o) v_1(t) \quad (3)$$

$$e(t_o) = v_2(t_o) - \hat{v}_2(t_o) \quad (4)$$

$$\hat{h}(t, t_o + 1) = \hat{h}(t, t_o) + \beta e(t_o) v_1(t), \quad \forall t \in U \quad (5)$$

where β and U denote the adaptation step-size and the support of the 1-D filter, respectively.

Consider the scenario where the desired response $v_2(t_o)$ is related to the input signal $v_1(t)$, through a system model $h(t, t_o)$, as $v_2(t_o) = \sum_t h(t, t_o) v_1(t)$. Under appropriate conditions, the adaptive filter coefficients $\hat{h}(t, t_o)$ closely approximate and track slow variations in $h(t, t_o)$. The adaptation step-size β determines the speed of convergence, tracking capability, and the closeness of the approximation [21, Ch. 5–6].

Based on the model of (2), an adaptive filtering framework can be readily extended to our image registration problem. For 2-D images, the adaptive filter takes the form of a 2-D finite impulse response (FIR) filter. Just as in the 1-D case, the adaptive filtering process consists of a prediction step and a filter adaptation step as shown in Fig. 1. Assuming that the LMS adaptation algorithm is used for the update, the process can be mathematically expressed as follows.

- 1) Filter output (Prediction step)

$$\hat{I}_2(x_o, y_o) = \sum_{(x, y \in R)} \hat{h}_b(x, y; x_o, y_o) I_1(x, y). \quad (6)$$

- 2) Estimation error

$$e(x_o, y_o) = I_2(x_o, y_o) - \hat{I}_2(x_o, y_o). \quad (7)$$

- 3) Filter adaptation (Update phase)

$$\begin{aligned} \hat{h}_a(x, y; x_o, y_o) &= \hat{h}_b(x, y; x_o, y_o) \\ &+ \beta e(x_o, y_o) I_1(x, y), \quad \forall (x, y) \in R. \end{aligned} \quad (8)$$

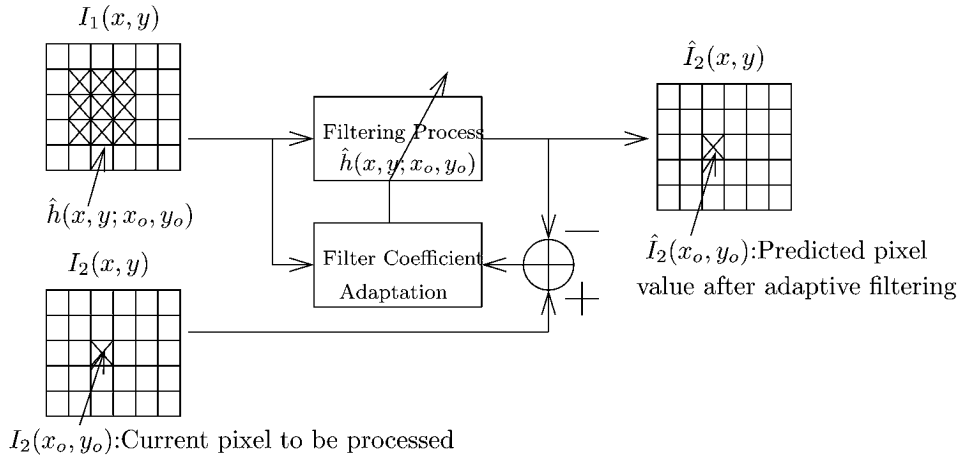


Fig. 1. Overview of 2-D adaptive filtering for images.

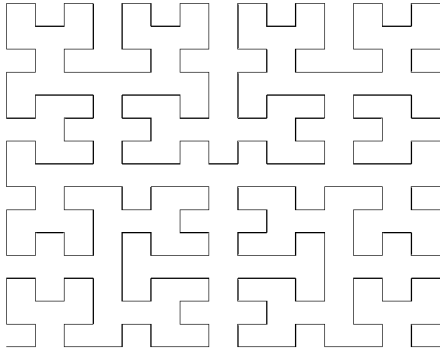


Fig. 2. Hilbert curve on a 16 × 16 square.

4) Initializing the filter for the next pixel, (x_n, y_n)

$$\hat{h}_b(x, y; x_n, y_n) = \hat{h}_a(x, y; x_o, y_o), \quad \forall (x, y) \in R \quad (9)$$

where $\hat{h}(\cdot)$ denotes the 2-D adaptive filter, R denotes the support of the filter, and β is the adaptation step-size. The subscripts b and a denote before and after adaptation, respectively.

Now, if the adaptive filter $\hat{h}(x, y; x_o, y_o)$ converges to the system model $h_o(x, y; x_o, y_o)$ of (2) and tracks changes in this system model, then the algorithm above provides us a mechanism for registering one image to the other. Though the mathematical formulation above bears a strong similarity with the 1-D adaptive filtering process, there are some key differences due to which the convergence and tracking behavior does not immediately follow in the 2-D case. We, therefore, consider this behavior next.

C. Contiguity Preservation Using Space-Filling Curves

The update step in the 1-D adaptive filter of (5) is along the natural sequence of temporal progression defined by the constraint of causality. For our formulation of the 2-D spatial adaptive filter; however, the scan order, i.e., the sequence in which

the image pixels are visited is not inherently defined. The scanning order plays a very important role in determining the behavior of the adaptive filter. Since the filter-adaptation is capable of tracking only “relatively slow” variations, its performance is best when smooth variations in the underlying system $h_o(x, y; x_o, y_o)$ appear smooth along the scan-order path. This is clearly not assured if the scan-order is discontinuous in the 2-D spatial domain. For this reason, the conventional raster scan order (moving from left to right and moving from top to bottom along successive horizontal lines of image pixels) is not appropriate for the filter adaptation. In addition to spatial continuity, it is also desirable (for the same reasons) to have a scan order that maps local neighborhoods in the 2-D spatial coordinates to localized regions in the scan sequence, i.e., to preserve spatial contiguity in the scan order (to the extent possible).

The desired property of contiguity preservation can be obtained in two dimensions through the use of suitable space-filling curves, which provide a mapping from multidimensional space into 1-D space. Because of its contiguity-preserving properties, Hilbert curves [24] are used to determine the scanning order of pixels. Fig. 2 shows a Hilbert curve for a 2-D region of size 16 × 16. As can be seen in the figure, traversal of the image pixels in the order indicated by this curve ensures that there are no jumps between spatially separated pixels.

D. Block-Based Adaptive Filtering

If unconstrained variations in filter coefficients on a pixel-by-pixel basis are allowed, the use of the current pixel information in the update step yields infinitely many solutions to the system identification problem, which can be overcome by regularization approaches. In the following, a regularization by block-based, as opposed to pixel-based, adaptation of the coefficients is proposed, whereby system response is assumed to remain constant over a block of pixels around the current pixel. To this effect, we first define a new optimization criterion, $E_{B_o}(x_o, y_o)$, which is the sum of estimation errors $e(x, y)$ over a block B_o of pixels

$$E_{B_o}(x_o, y_o) = \sum_{(x_b, y_b) \in B_o} e^2(x_b, y_b) \quad (10)$$

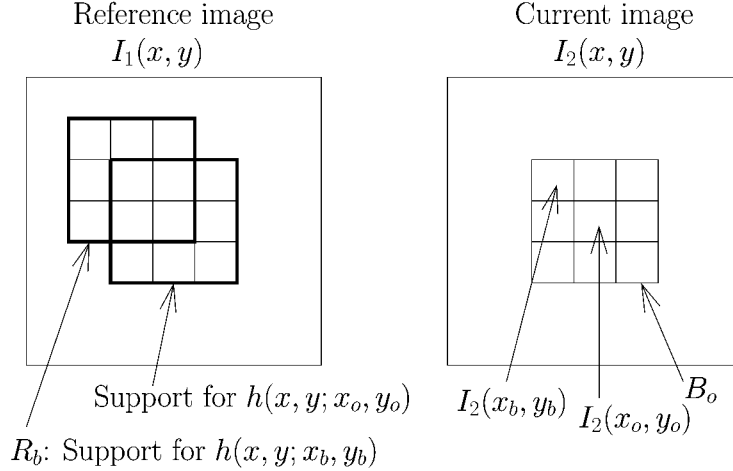


Fig. 3. Block-based adaptive filtering: On the reference image, $I_1(x, y)$ filter support is shown for the adaptive filters that correspond to pixels within the block, B_o , in the current image, $I_2(x, y)$.

where B_o is the block around the current pixel, $I_2(x_o, y_o)$, as shown in Fig. 3. When we integrate (6) and (7) with (10), we obtain

$$\begin{aligned}
 E_{B_o}(x_o, y_o) &= \sum_{(x_b, y_b) \in B_o} (I_2(x_b, y_b) - \hat{I}_2(x_b, y_b))^2 \\
 E_{B_o}(x_o, y_o) &= \sum_{(x_b, y_b) \in B_o} \left(I_2(x_b, y_b) \right. \\
 &\quad \left. - \sum_{(x_i, y_i) \in R_b} \hat{h}_b(x_i, y_i; x_b, y_b) I_1(x_i, y_i) \right)^2
 \end{aligned} \quad (11)$$

where R_b is the support of the adaptive filter for the pixel, $I_2(x_b, y_b)$, defined on the first image, I_1 , as shown in Fig. 3. As a regularization constraint, we impose the uniform system response (i.e., $\hat{h}_b(x_i, y_i; x_b, y_b) = \hat{h}_b(x_i, y_i; x_o, y_o)$, $\forall (x_b, y_b) \in B_o$) into (12), and obtain the following equation for $E_{B_o}(x_o, y_o)$

$$\begin{aligned}
 E_{B_o}(x_o, y_o) &= \sum_{(x_b, y_b) \in B_o} \left(I_2(x_b, y_b) \right. \\
 &\quad \left. - \sum_{(x_i, y_i) \in R_b} \hat{h}_b(x_i, y_i; x_o, y_o) I_1(x_i, y_i) \right)^2.
 \end{aligned} \quad (13)$$

The updated adaptive filter coefficients $\hat{h}_a(x_i, y_i; x_o, y_o)$ are then computed by taking the derivative of $E_{B_o}(x_o, y_o)$ with respect to the filter coefficients, and using the derivative in the least mean square formulation. Equation (14) denotes the updated filter coefficients, $\hat{h}_a(x_i, y_i; x_o, y_o)$, which is very similar to (8), except that a block of pixels over which the system response is uniform for each pixel within the block, B_o [see (14), shown at the bottom of the page] where (x_{ib}, y_{ib}) denotes the spatial location where the adaptive filter is convolved with the reference image, $I_1(x, y)$, for the corresponding pixel, $I_2(x_b, y_b) \in B_o$.

E. Implementation Issues

The implementation of the algorithm requires a number of additional choices. To prevent the amplification of gradient noise and to increase the convergence rate, the normalized LMS adaptation algorithm [21, Ch. 6] is used. This results in a 2-D adaptive filter closer to the system model, and, therefore, better image registration.

The choice of adaptation step-size, β , is crucial in the 2-D LMS adaptation algorithm. This is because the adaptive filter coefficients $\hat{h}(x, y; x_o, y_o)$ are updated at a speed that is determined by β . If the variation in the system model $h_o(x, y; x_o, y_o)$ from one pixel to the next one is significant, then filter adaptation from one pixel to the next one through LMS may not be possible unless the right adaptation step-size is chosen. On the other hand, if the chosen β is too large, then the optimum filter coefficients may never be achieved due to gradient estimation noise [21, Ch. 5].

While scanning the images with a Hilbert curve, the input images are divided into overlapping squares, because the Hilbert

$$\begin{aligned}
 \hat{h}_a(x_i, y_i; x_o, y_o) &= \hat{h}_b(x_i, y_i; x_o, y_o) \\
 &\quad + \beta \sum_{(x_b, y_b) \in B_o} \left(I_2(x_b, y_b) - \sum_{(x_i, y_i) \in R_b} \hat{h}_b(x_i, y_i; x_o, y_o) I_1(x_i, y_i) \right) \\
 &\quad \times I_1(x_{ib}, y_{ib})
 \end{aligned} \quad (14)$$

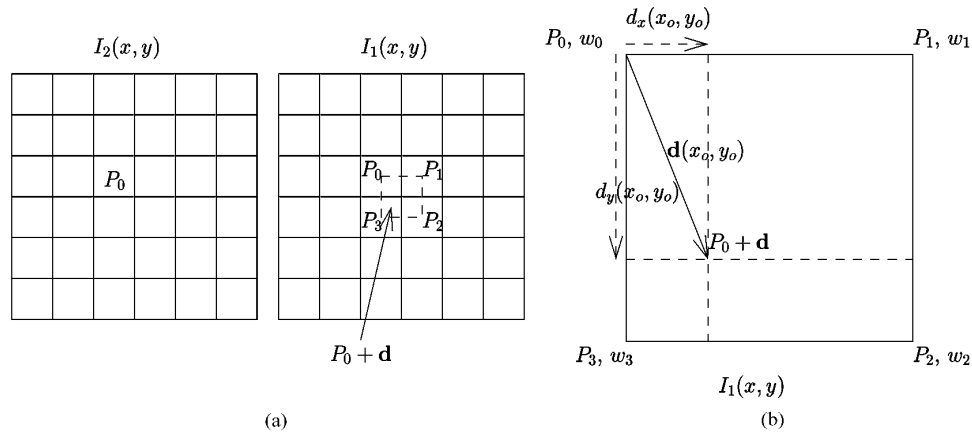


Fig. 4. Motion modeling through adaptive filtering. (a) $P_0 = (x_o, y_o)$ in the second image I_2 corresponds to $P_0 + \mathbf{d} = ((x_o + d_x), (y_o + d_y))$ in the first image I_1 . (b) Corresponding weights for the bi-linearly interpolation of gray-level value at $P_0 + \mathbf{d}$ in I_1 .

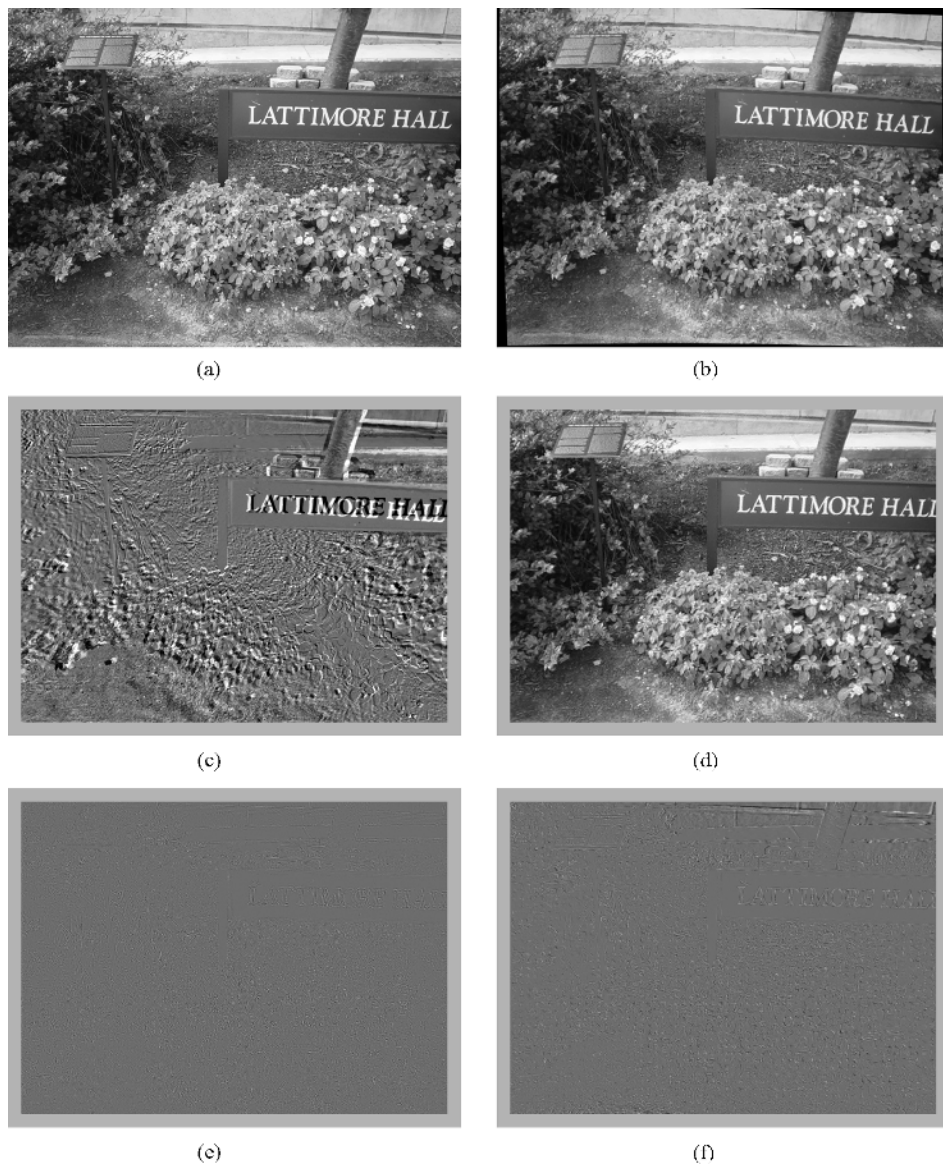


Fig. 5. Registration example under Stirmark attack. (a) Original flower image. (b) Distorted image after applying Stirmark. (c) Difference image between (a) and (b). (d) Registered image with the proposed registration technique. (e) Difference image between (a) and (d). (f) Difference image by applying the registration algorithm in [17].

curve requires the 2-D region to be a square. A Hilbert curve is computed for each square, and the proposed 2-D adaptive fil-

tering technique is performed on each square consecutively. The estimated adaptive filter on one square is used as the initial adap-

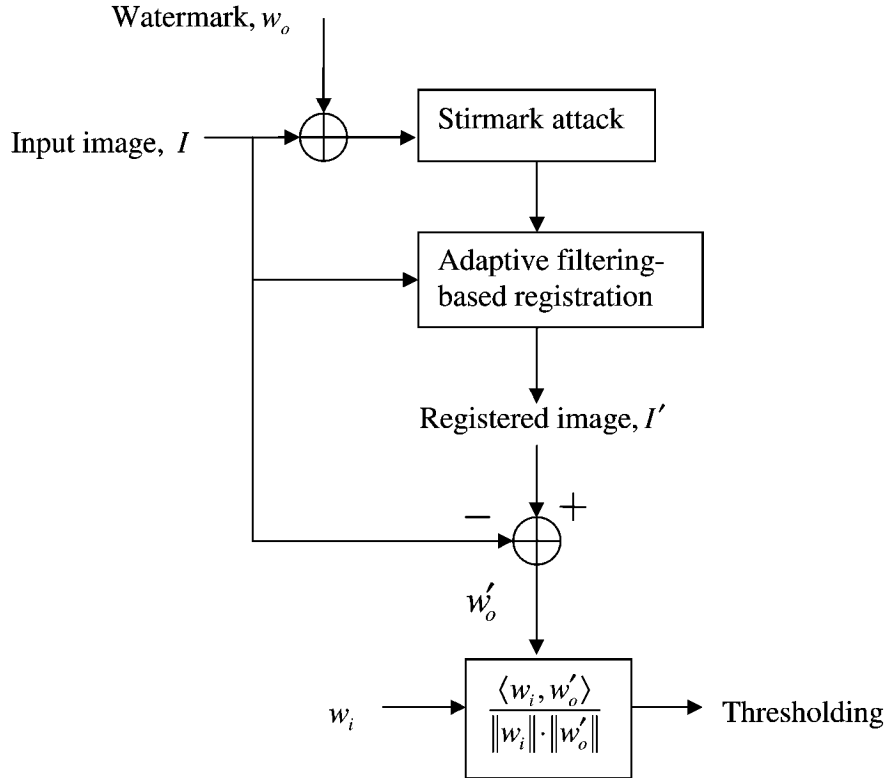


Fig. 6. Watermark detection setup for registration evaluation on Stirmark attacked images.

tive filter on the next square in such a way that the scanning order of the whole input image is guaranteed to be contiguity preserving. Since the adaptive filter requires a number of iterations (pixel-visits) for the initial convergence, the initial region (in the scan-order) may show a relatively large registration error. This is, however, readily handled by revisiting this region in a second pass originating from a region in which convergence has been achieved (for instance, by reversing the scan-order for the small initial region).

The “size” of the adaptive filter should be large enough to represent the spatially varying system model $h_o(x, y; x_o, y_o)$ at each point. This clearly requires a rather large filter in situations where there are large displacements between the images. A simple increase in filter size would adversely affect the convergence behavior in addition to the increased computational burden for adaptation. We, therefore, keep track of large shifts independently as integer pixel shifts in x and y . This is achieved by adjusting these values at each pixel to ensure that the central peak of the adaptive filter coefficients is located near the center of the filter support.

Since the proposed registration technique handles only local misregistration, in scenarios where the images exhibit large global misregistration, we use an initialization step that positions the support of the adaptive filter in the appropriate neighborhood. This is achieved by setting the initial values of integer pixel shifts to a crude estimate of the displacement for the first pixel in the scan order (with filter coefficients set to an impulse). Parametric models for global registration (e.g., affine/pseudoperspective [2]) may be utilized for this purpose.

Model parameters may be estimated using feature correspondences between images [25] or using hierarchical estimation methods [11].

III. MODEL VALIDATION: CONFORMANCE TO A LOCAL DISPLACEMENT FIELD

In this section, we discuss the conditions under which the adaptive filter coefficients in (2) conform to an underlying local displacement field $\mathbf{d}(x, y) = (d_x(x, y), d_y(x, y))$ in (1).

If, for the moment, we assume that the displacement vector $\mathbf{d}(x_o, y_o)$ at pixel (x_o, y_o) is integer valued, i.e., does not have a fractional part, then the coefficients of the adaptive filter $h_o(x, y; x_o, y_o)$ should converge to an impulse at $(x_o + d_x(x_o, y_o), y_o + d_y(x_o, y_o))$.

In the more realistic case, when the displacement vector $\mathbf{d}(x, y)$ is subpixel valued, i.e., it includes both an integer and a fractional part, the adaptive filter $h_o(x, y; x_o, y_o)$ should take the form of an interpolation filter. More specifically, let us assume that there is subpixel displacement $(d_x(x_o, y_o), d_y(x_o, y_o))$ where

$$I_2(x_o, y_o) = I_1(x_o + d_x(x_o, y_o), y_o + d_y(x_o, y_o)). \quad (15)$$

Fig. 4(a) denotes the spatial location of a point $P_0 = (x_o, y_o)$ in the second image I_2 and its corresponding location $P_0 + \mathbf{d}$ in the first image I_1 . As shown in Fig. 4, the closest (integer) pixel locations to $P_0 + \mathbf{d}$ in the first image I_1 are denoted as $P_0, P_1, P_2,$ and P_3 .

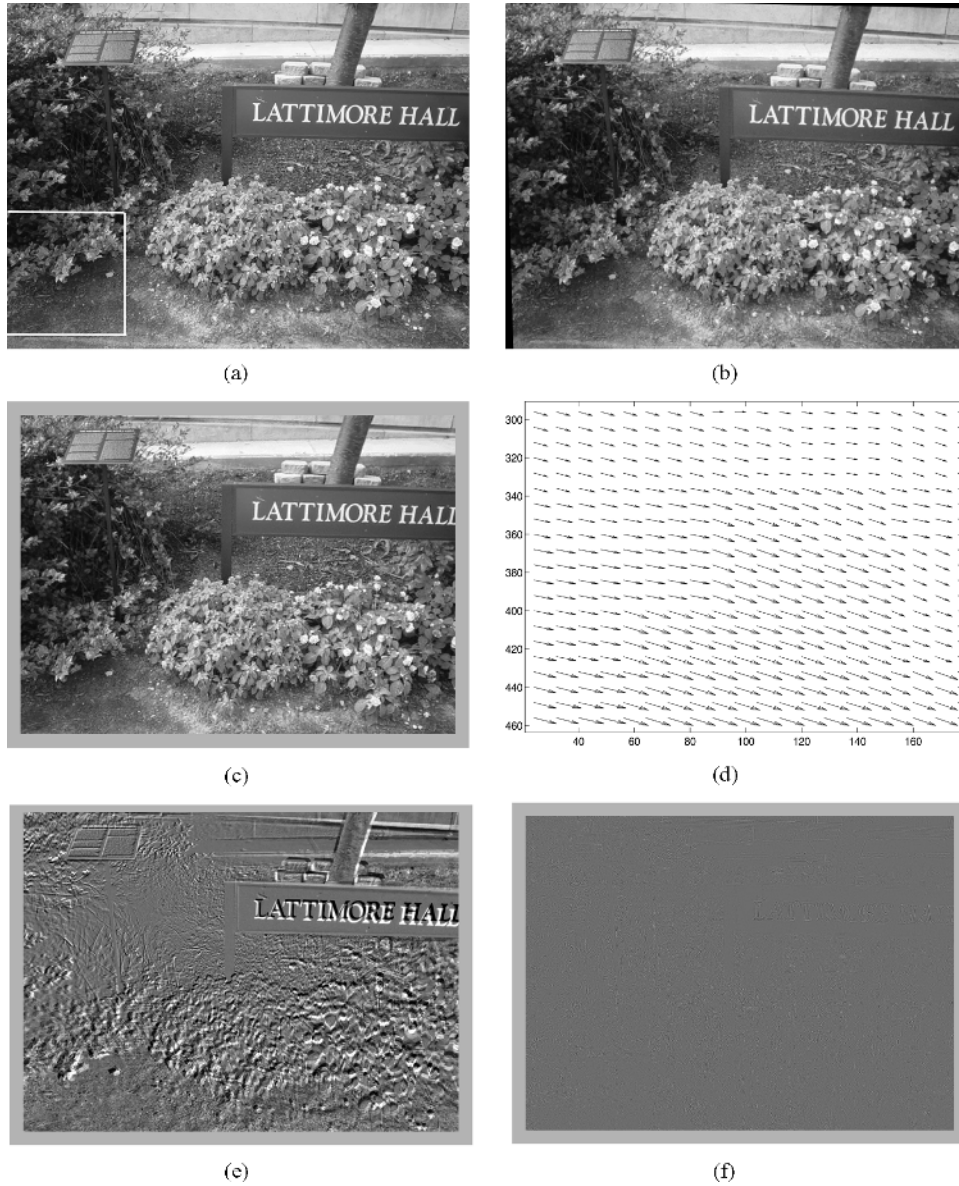


Fig. 7. Watermark detection example. (a) Undistorted watermark image. (b) Distorted watermark image after applying Stirmark with a bending factor of 0.3. (c) Registered watermark image. (d) Estimated motion field in the highlighted area in (a). (e) Difference image between (a) and (b). (f) Difference image after applying the proposed registration technique.

Using the displacement vector, $\mathbf{d}(x_o, y_o)$, we can compute the gray-level value of point $P_0 + \mathbf{d}$ in the first image, I_1 by some finite impulse response (FIR) interpolation filter, e.g., the bi-linear interpolation. Fig. 4(b) denotes the weights (w_0, w_1, w_2, w_3) for the corresponding points, (P_0, P_1, P_2, P_3) , that are used in bi-linear interpolation. These weights can be expressed in terms of the displacement vector $\mathbf{d}(x_o, y_o) = (d_x(x_o, y_o), d_y(x_o, y_o))$ as

$$\begin{aligned}
 w_0 &= (1 - d_x(x_o, y_o))(1 - d_y(x_o, y_o)) \\
 w_1 &= d_x(x_o, y_o)(1 - d_y(x_o, y_o)) \\
 w_2 &= d_x(x_o, y_o)d_y(x_o, y_o) \\
 w_3 &= (1 - d_x(x_o, y_o))d_y(x_o, y_o).
 \end{aligned} \tag{16}$$

Then

$$I_2(P_0) = w_0I_1(P_0) + w_1I_1(P_1) + w_2I_1(P_2) + w_3I_1(P_3). \tag{17}$$

Comparing (17) with (2), we expect that the adaptive filter $h_o(x, y; x_o, y_o)$ should have a small number of significant (nonzero) coefficients peaked around $(x_o + d_x(x_o, y_o), y_o + d_y(x_o, y_o))$.

From these observations, we can state the following two criteria to validate whether the adaptive filter coefficients conform to an underlying local displacement vector between the images at each pixel (x_o, y_o) : 1) we expect the filter to have a single peak within its region of support, 2) if the illumination/exposure differences between the images are negligible, the filter coefficients should sum up to unity. We use empirically determined thresholds to determine if the first criterion holds, assuming a

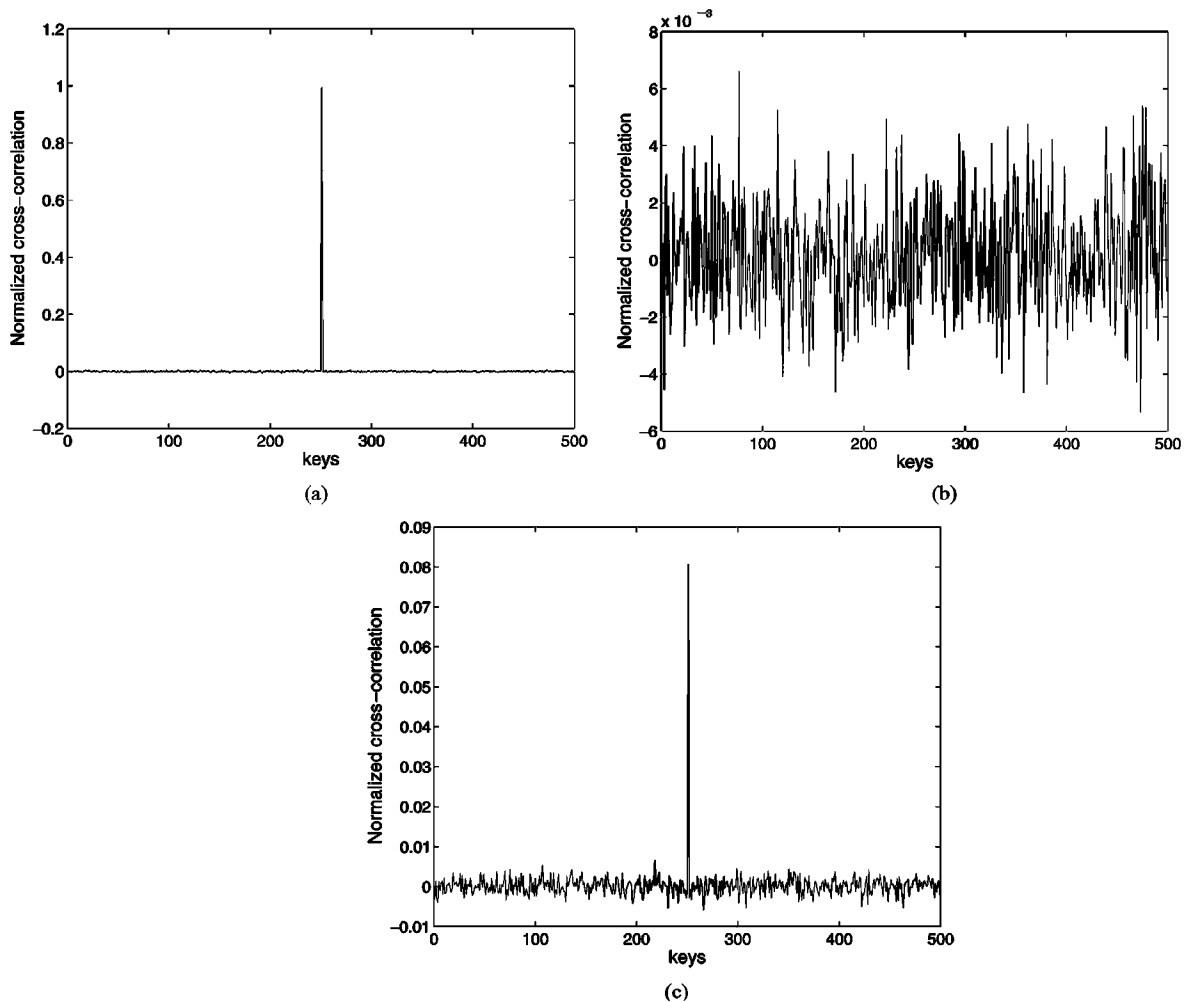


Fig. 8. Detector response of the (a) undistorted watermark image, (b) distorted watermark image and (c) registered watermark image.

single peak if all other local maxima were at least 50% below the global maximum.

In our experiments, we observe that the criteria are not satisfied at some pixels under the following situations: a) when the image region about (x_o, y_o) lacks sufficient texture (the aperture problem), b) if there is occlusion around (x_o, y_o) , or c) when there are abrupt changes in the local motion field about (x_o, y_o) .

IV. APPLICATIONS AND EXPERIMENTAL RESULTS

A. Compensation of Random Geometric Distortions (Stirmark)

The effectiveness of the proposed registration scheme is tested under Stirmark attack [9] by performing two sets of experiments. In the first experiment, performance of the proposed technique is compared with the registration algorithm given in [17], which is designed to register geometrically distorted images for watermark recovery. In the second experiment, performance of the proposed registration scheme is measured in terms of watermark detection.

In the experiments, an 8-bit gray scale image of size 640×480 pixels is used as the original image. The following parameters are utilized in the proposed adaptive filtering-based image registration algorithm: The size of the adaptive filter is

set to 9, i.e., a 9×9 filter is chosen. The adaptation step-size and scanning curve are respectively 0.2, and a Hilbert curve. As a regularization constraint, adaptive filter coefficients are assumed to be constant over a 3×3 window during the adaptation process. Fig. 5 presents the results for the first experiment, where performance of the proposed registration technique is compared with the image registration algorithm given in [17]. The resulting pSNR values for the difference images given in Fig. 5(e) and (f) are 28.12 and 28.35 dB, respectively. Although pSNR values of the difference images obtained by the proposed registration scheme and the registration algorithm presented in [17] are very close, visual observation of the difference images reveals that our proposed scheme results in a registered image of higher quality.

In the second set of experiments, we evaluate the watermark detection performance with the proposed registration technique. The original image I used in the experiment is shown in Fig. 5(a). In this image, we embed a spread spectrum watermark consisting of a pseudorandom noise sequence, w_o , that is uniformly distributed between integers -5 and 5 . The seed for the watermark generator serves as the unique watermark key. As a desynchronization attack, the watermarked image is then geometrically distorted using Stirmark (a bending factor

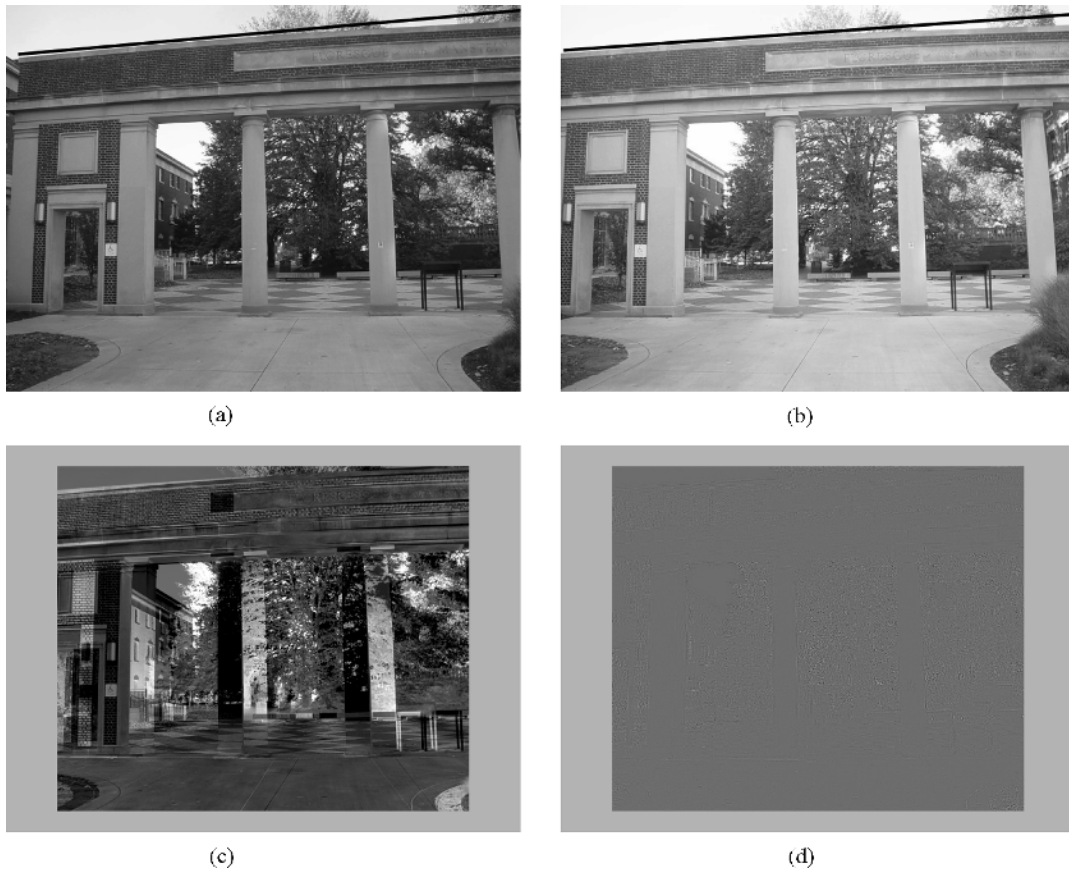


Fig. 9. Registration example with radial distortion and projective motion. A straight line is drawn on top of both images to show the radial distortion. (a) Reference input image. (b) Current input image. (c) Difference image between (a) and (b). (d) Difference image after applying the proposed registration technique.

of 0.3, which is consistent with real-life attacks is utilized). The resulting distorted image is then registered to the original image using the proposed registration technique and the watermark is extracted from the registered image, I' using the original image, I in a nonblind detection scenario by correlating the extracted watermark w'_o obtained as the difference of the original and registered images with a pseudorandom sequence generated with the watermark key (for the correct key, this is identical to the embedded watermark w_i). Fig. 6 summarizes the experimental setup schematically.

Fig. 7(a) and (b) depicts the watermarked image and distorted watermarked image, respectively. After applying the proposed registration technique with the input parameters given at the start of this section, the registered image, I' shown in Fig. 7(c) is obtained. Fig. 7(e) and (f) shows initial difference image and the resulting difference image, respectively. The pSNR values of the difference images are 16.11 and 29.46 dB for Fig. 7(e) and (f), respectively.

We evaluate the efficacy of the registration process for recovering from geometric distortion attacks by considering the difference between the response to the correct watermark key and to other random watermark keys. A clear separation between these is preferable in order to obtain good probability of detection at a low false alarm rate [26]. Fig. 8 shows the watermark detection results for the undistorted watermark image, Fig. 7(a), distorted watermark image, Fig. 7(b) and the registered image, Fig. 7(c). In each case, the evaluation considers the watermark

detection process utilizing the correct key and 499 other random keys. As shown in Fig. 8(a), when there is no Stirmark attack and registration, correlation of the original watermark with itself is very close to 1, which is 0.995, as expected and the response to random keys is close to zero. When there is Stirmark attack and no registration, watermark is extracted from distorted watermark image. Fig. 8(b) shows that correlation of the extracted watermark with pseudorandom sequences with random watermark keys is statistically indistinguishable from the response to the correct key (indicating that the attack would be effective). Fig. 8(c) shows the correlation detector responses after the distorted watermark image is registered using the proposed registration scheme. We see that in this case, the correlation of the extracted watermark with the pseudorandom sequence generated with the original watermark key is almost 12 times higher than the correlation with pseudorandom sequences generated with arbitrary watermark keys. Therefore, we can conclude that under Stirmark attack, our proposed registration scheme provides a very high quality registered image, and helps to detect the original watermark in a reliable way.

B. Compensation of Camera Lens Distortions

In this section, we test the performance of the proposed registration scheme under another locally varying spatial geometric distortion, i.e., “radial distortion.” Due to manufacturing limitations and imperfections, all imaging devices introduce

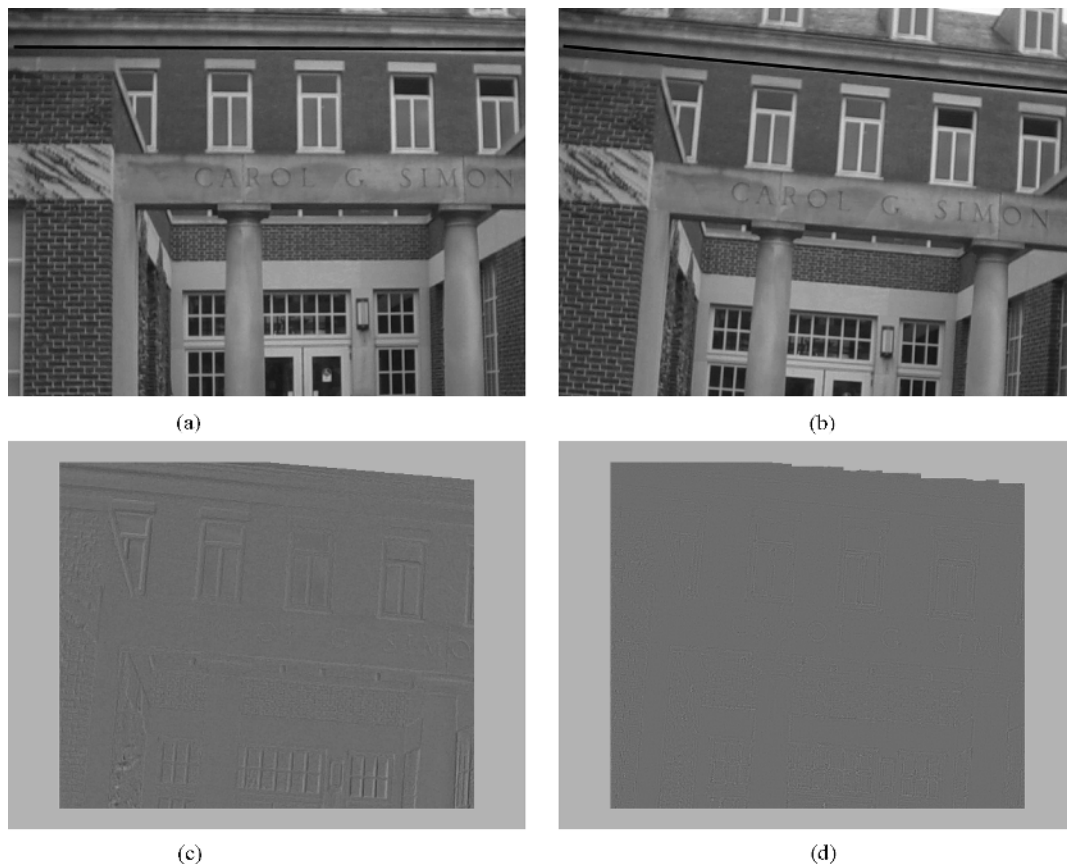


Fig. 10. Registration example with radial distortion and affine motion. A straight line is drawn on top of both images to show the radial distortion. (a) Reference input image. (b) Current input image. (c) Difference image with feature-based global alignment. (d) Difference image by applying the proposed registration technique.

some nonlinear distortion in captured images. Radial distortion is one of the most severe nonlinear distortions observed, wherein image points are displaced with respect to their ideal positions—inward or outward along the radial direction from a center [27]. If the radial displacement is outward, then the distortion is called pincushion. Otherwise, it is called barrel distortion.

We present here two experiments to demonstrate the performance of the proposed local registration technique in the presence of radial distortion. In the first experiment, we captured two images by panning a low-end digital camera mounted on a tripod. Both images contain barrel distortion as shown in Fig. 9(a) and (b). In the second experiment, we captured two other images from a different scene by changing roll angle of the same camera, i.e., by rotating the camera in the image plane. Fig. 10(a) and (b) shows the input images captured for the second experiment, which contain less amount of radial distortion than the ones used in the first experiment, though the amount of camera motion is larger in this case.

In both experiments, we utilize the following parameters in the adaptive-filtering based registration technique. Size of the adaptive filter is set to 13×13 . The adaptation step-size and block size for the block-LMS are respectively, 0.2 and (3×3) . Hilbert curve is utilized to scan the input image as explained in Section II-E. For the first experiment, Fig. 9(c) and (d) show the difference image between the input images,

and the resulting difference image after applying the proposed registration method, respectively. The pSNR of the resulting error image using the proposed technique is 31.63 dB. When c-scan (L-R on row 1, R-L on row 2, L-R on row 3, etc.) is used as the image scan order instead of Hilbert curves in the adaptive filtering framework, pSNR of the error image decreases to 28.59 dB. As observed from Fig. 9(d), the registration error is concentrated mainly in the regions of the tree leaves. This arises primarily because of independent motion of the leaves between the exposures.

In the second experiment, in order to handle the larger camera rotation, suitable initialization of the adaptive filter support is required. For this purpose, a pseudoperspective motion model (i.e., eight parametric model) is estimated using image feature point correspondences that are determined using a feature detection and matching algorithm [25]. The estimated motion model parameters are used to initialize the adaptive filter support as described in Section II-E. Subsequently, the adaptive filter is updated by applying the proposed registration technique following the image scan order. Fig. 10(d) shows the resulting difference image by applying the proposed registration method. Fig. 10(c) shows the difference image obtained using the (global) pseudoperspective motion model.

The pSNR values for the error images obtained with the global pseudoperspective model and the proposed technique are 34.0 and 36.53 dB, respectively. This illustrates the advantage

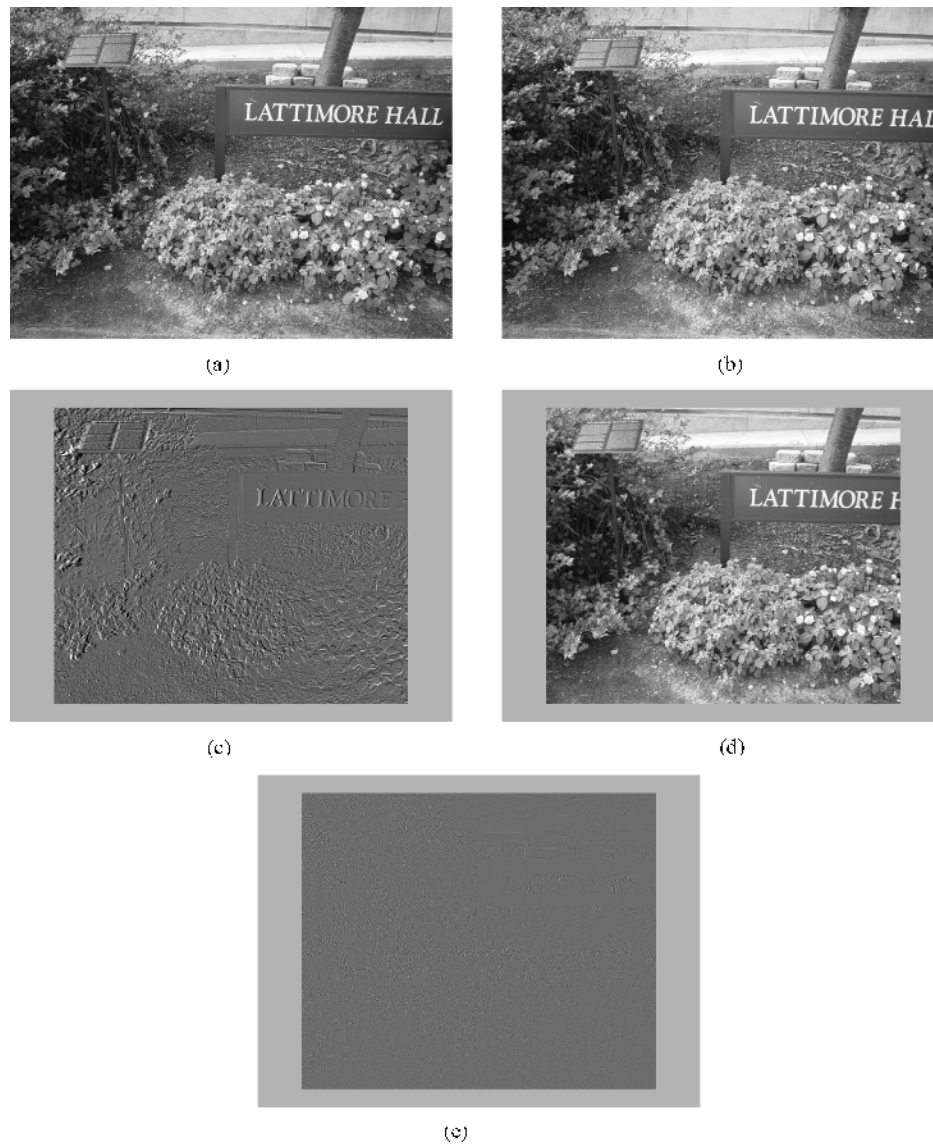


Fig. 11. Registration example of multiview images with nonparametric local motion. (a) Reference input image. (b) Current input image. (c) Difference image with gradient-based global alignment. (d) Registered current image with the proposed registration technique. (e) Difference image by applying the proposed registration technique.

of the proposed method. Furthermore, when c-scan is used as the image scan order in the adaptive filtering framework, pSNR of the resulting error image decreases to 34.5 dB, which shows the improvement provided by the Hilbert curves.

C. Registration of Multiview Images With Nonparametric Local Motion

In this section, we evaluate performance of the proposed local image registration technique in a multiview capture scenario where image motion cannot be modeled by a parametric motion model. Fig. 11(a) and (b) shows two input images that are captured by a moving digital camera. The scene consists of multiple objects at different depths and is, therefore, nonplanar. Consequently, parametric global registration models intended for planar scenes are inadequate.

We initialize the support of the adaptive filter (see Section II-E) at the first pixel in the scan order using a hierarchical

parametric registration algorithm [11].¹ The adaptive filter is then updated following the scan order, using the proposed registration technique. Fig. 11(d) shows the resulting registered image using the following set of parameters in the adaptive filtering framework: 1) adaptation step size is set to 0.2, 2) a constant size adaptive filter (13×13) is used in the adaptation process, 3) block size for the block LMS is set to (3×3) , and 4) Hilbert Curve is utilized to scan the image.

When the parametric registration algorithm [11] is applied globally, the registration error image shown in Fig. 11(c) is obtained. The pSNRs of the error images with global alignment and the proposed registration method are 20.60 and 29.40 dB, respectively, which demonstrate the improvement with the proposed method. For comparison, the proposed registration technique is also applied using a c-scan order, which produces a

¹Due to the highly textured image content, feature matching performs poorly here so estimation from feature correspondences is not employed.

pSNR of 27.39 dB. We also compared the proposed technique with the hierarchical Lucas–Kanade technique [10]. For this case, the best pSNR obtained for the registered image is 25.57 dB, which is obtained with a window size of 9×9 (i.e., number of pyramid levels is 3).

We observe that the filter coefficients do not satisfy the “single peak” criterion (see Section III), in regions where there are occlusions. Because this occurs only in a small number of pixels, the visual and PSNR performance are not significantly affected.

V. CONCLUSION

In this paper, we propose a new local image registration technique based on an adaptive filtering framework that can handle smooth spatial variations in registration. We show the effectiveness of the technique both visually and numerically in a number of application scenarios. Experimental results demonstrate that the proposed technique can be used to register images perturbed using Stirmark (for the purpose of watermark recovery), images with radial lens distortion, and multiview images with small camera motion where parametric models are inadequate. In these applications, the technique offers significant improvements over global registration alone and over other common local registration techniques.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their detailed comments and suggestions on an earlier version of this paper; they have significantly helped to improve the quality of the presentation and helped weed out misstatements and ambiguities.

REFERENCES

- [1] B. Zitova and J. Flusser, “Image registration methods: A survey,” *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Nov. 2003.
- [2] 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.
- [3] H. Shum and R. Szeliski, “Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment,” *Int. J. Comput. Vis.*, vol. 36, pp. 101–130, 2000.
- [4] Y. Caspi and M. Irani, “Spatio-temporal alignment of sequences,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 11, pp. 1409–1424, Nov. 2002.
- [5] H. Sawhney and R. Kumar, “True multi-image alignment and its application to mosaicing and lens distortion correction,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 3, pp. 235–243, Mar. 1999.
- [6] M. Irani and P. Anandan, “All about direct methods,” in *Proc. Int. Workshop Vision Algorithms*, 1999, pp. 267–277.
- [7] P. H. S. Torr and A. Zisserman, “Feature based methods for structure and motion estimation,” in *Proc. Int. Workshop Vision Algorithms*, 1999, pp. 278–295.
- [8] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. Assoc. Comput. Mach.*, vol. 24, no. 6, pp. 381–395, 1981.

- [9] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Attacks on copyright marking systems,” in *Proc. 2nd Int. Information Hiding*, Portland, OR, Apr. 1998, pp. 219–239.
- [10] B. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” presented at the DARPA Image Understanding Workshop 1981.
- [11] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, “Hierarchical model-based motion estimation,” in *Proc. Eur. Conf. Computer Vision*, May 1992, pp. 237–252.
- [12] H. Liu, R. Chellappa, and A. Rosenfeld, “Accurate dense optical flow estimation using adaptive structure tensors and a parametric model,” *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1170–1180, Oct. 2003.
- [13] A. M. Tekalp, P. J. L. Van Beek, C. Toklu, and B. Gunsel, “2D mesh-based visual object representation for interactive synthetic/natural video,” *Proc. IEEE*, vol. 86, no. 6, pp. 1029–1051, Jun. 1998.
- [14] H. Sawhney, “Simplifying motion and structure analysis using planar parallax and image warping,” in *Proc. Int. Conf. Pattern Recognition*, 1994, pp. A403–A408.
- [15] C. Toklu, A. T. Erdem, M. I. Sezan, and A. M. Tekalp, “Tracking motion and intensity-variations using hierarchical 2-D mesh modeling for synthetic object transfiguration,” *CVGIP: Graph. Models Image Process.*, vol. 58, no. 6, pp. 553–573, Nov. 1996.
- [16] M. Irani, P. Anandan, and M. Cohen, “Direct recovery of planar-parallax from multiple frames,” presented at the Int. Conf. Computer Vision Workshop: Vision Algorithms Sep. 1999.
- [17] P. Loo and N. G. Kingsbury, P. W. Wong and E. J. Delp, Eds., “Motion estimation based registration of geometrically distorted images for watermark recovery,” in *Proc. SPIE: Security and Watermarking of Multimedia Contents III*, Jan. 2001, vol. 4314, pp. 606–617.
- [18] P. J. O. Doets, “Complexity-Scalable Compensation of Geometrical Distortions in Image Watermarking,” M.Sc. thesis, Dept. Mediamatics, Delft Univ. Technol., Delft, The Netherlands, 2003.
- [19] J. Du, C. Woo, and B. Pham, “Recovery of watermark using differential affine motion estimation,” presented at the Austral. Information Security Workshop 2005.
- [20] F. Šroubek and J. Flusser, “Multichannel blind deconvolution of spatially misaligned images,” *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 874–883, Jul., 2005.
- [21] S. Haykin, *Adaptive Filter Theory*, Fourth ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [22] F. A. P. Petitcolas, “Watermarking schemes evaluation,” *IEEE Trans. Signal Process.*, vol. 17, no. 1, pp. 58–64, Jan. 2000.
- [23] G. Sharma, “Show-through cancellation in scans of duplex printed documents,” *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 736–754, May 2001.
- [24] H. Sagan, *Space-Filling Curves*. Berlin, Germany: Springer, 1994.
- [25] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [26] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*. San Francisco, CA: Morgan Kaufmann, 2002.
- [27] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Englewood Cliffs, NJ: Cambridge Univ. Press, 2004.



Gulcin Caner (S'00) received the B.S. degree (with high honors) in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2000, and the M.S. degree in electrical and computer engineering from the University of Rochester, Rochester, NY, in 2002. She is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Rochester.

Her research interests lie in the area of image and video processing, specifically image registration and its applications in the areas of image enhancement and mosaicking.



A. Murat Tekalp (S'80–M'84–SM'91–F'03) received the M.S. and Ph.D. degrees in electrical, computer, and systems engineering from Rensselaer Polytechnic Institute (RPI), Troy, NY, in 1982 and 1984, respectively.

From December 1984 to August 1987, he was with Eastman Kodak Company, Rochester, NY. He joined the Electrical and Computer Engineering Department at the University of Rochester in September 1987, where he is currently a Distinguished Professor. Since June 2001, he also has been with Koc

University, Istanbul, Turkey. His research interests are in the areas of digital image and video processing, including video compression and streaming, video filtering for high-resolution, video segmentation, object tracking, content-based video analysis and summarization, multicamera surveillance video processing, and protection of digital content. He authored the book *Digital Video Processing* (Prentice-Hall, 1995). He holds five U.S. patents. His group contributed technology to the ISO/IEC MPEG-4 and MPEG-7 standards.

Dr. Tekalp was named a Distinguished Lecturer by the IEEE Signal Processing Society in 1998. He has chaired the IEEE Signal Processing Society Technical Committee on Image and Multidimensional Signal Processing (January 1996 to December 1997). He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1990 to 1992), IEEE TRANSACTIONS ON IMAGE PROCESSING (1994 to 1996), and the *Journal of Multidimensional Systems and Signal Processing* (1994 to 2002). He was an area editor for the *Journal of Graphical Models and Image Processing* (1995 to 1998). He was also on the editorial board of the *Journal of Visual Communication and Image Representation* (1995 to 2002). He was appointed as the Technical Program Chair for the 1991 IEEE Signal Processing Society Workshop on Image and Multidimensional Signal Processing, the Special Sessions Chair for the 1995 IEEE International Conference on Image Processing, the Technical Program Co-Chair for IEEE ICASSP 2000, Istanbul, Turkey, and the General Chair of IEEE International Conference on Image Processing (ICIP), Rochester, in 2002. He is the Founder and first Chairman of the Rochester Chapter of the IEEE Signal Processing Society. He was elected as the Chair of the Rochester Section of IEEE from 1994 to 1995. At present, he is the Editor-in-Chief of the EURASIP journal *Signal Processing: Image Communication*. He serves as the Chairman of the Electronics and Information Technology Group of the Turkish Science and Technology Foundation (TUBITAK). He also serves as an independent expert to review projects for the European Commission.



Gaurav Sharma (S'88–M'96–SM'00) received the B.E. degree in electronics and communication engineering from Indian Institute of Technology Roorkee (formerly the University of Roorkee), India, in 1990, the M.E. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, in 1992, and the M.S. degree in applied mathematics and the Ph.D. degree in electrical and computer engineering from North Carolina State University (NCSSU), Raleigh, in 1995 and 1996, respectively.

From August 1992 to August 1996, he was a Research Assistant at the Center for Advanced Computing and Communications, Electrical and Computer Engineering Department, NCSU. From August 1996 to August 2003, he was with Xerox Research and Technology, Webster, NY, initially as a Member of Research Staff and subsequently in the position of Principal Scientist. Since Fall 2003, he has been an Associate Professor with the University of Rochester, Rochester, NY. His research interests include multimedia security and watermarking, color science and imaging, signal restoration, and halftoning.

Dr. Sharma is a member of Sigma Xi, Phi Kappa Phi, Pi Mu Epsilon, IS&T, and the IEEE Signal Processing Society. He was the 2003 Chair for the Rochester Chapter of the IEEE Signal Processing Society and is the Treasurer for the Rochester Section. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and the *Journal of Electronic Imaging*.



Wendi Heinzelman (S'95–M'00–SM'05) received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, in 1995, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1997 and 2000, respectively.

She is an Assistant Professor in the Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY. Her current research interests lie in the area of wireless communications and

networking, mobile computing, and multimedia communication.

Dr. Heinzelman is a member of Sigma X and the ACM. She received the National Science Foundation CAREER award in 2005 for her work on cross-layer architectures for wireless sensor networks. She also received the ONR Young Investigator Award in 2005 for her work on balancing resource utilization in wireless sensor networks.