AN ADAPTIVE FILTERING FRAMEWORK FOR IMAGE REGISTRATION

Gulcin Caner[a], A. Murat Tekalp[a,b], Gaurav Sharma[a], Wendi Heinzelman[a]

[a] Electrical and Computer Engineering Dept., University of Rochester, Rochester, NY, 14627-0126
[b] College of Engineering, Koc University, Istanbul, Turkey
{caner, tekalp, gsharma, wheinzel}@ece.rochester.edu

ABSTRACT

Image registration is a fundamental task in both image processing and computer vision. Here, we present a novel method for local image registration based on adaptive filtering techniques. We utilize an adaptive filter to estimate and track correspondences among multiple images containing overlapping views of common scene regions. Image pixels are traversed in an order established by space-filling curves, to preserve the contiguity and hence track locally varying registration changes. The algorithm differs from pre-existing work on image registration in that it requires only local information and relatively low computational effort. These characteristics render the method suitable for deployment in imaging sensor networks, toward which the current work is directed. We evaluate the performance of the proposed algorithm using images captured with a digital camera in various real-world scenarios. Experimental results show that the proposed method can significantly improve accuracy and robustness over a global 2-D parametric registration and can also outperform the local registration algorithm based on the Lucas-Kanade [1] optical flow technique.

1. INTRODUCTION

Image registration is needed for various applications, such as noise suppression, mosaicking, super-resolution, object tracking, and 3D scene reconstruction. Additionally, motion estimation among video frames can also be considered an instance of image registration. A large number of techniques have been developed to solve different variants of this problem, in both image processing and computer vision areas.

The techniques fall in two main classes: a) methods that rely on only the image data and make no assumptions about the underlying camera or scene geometry and b) techniques that are designed specifically for images of 3-D scenes that assume an underlying scene and camera model (either known or unknown). Predominant in the latter category are methods from computer vision, where image registration is usually performed as a prelude to computing 3-D scene structure, and often underlying 3-D camera and scene geometry is used as the basis of registration algorithms [2]. In addition, applications for video have inspired a number of motion estimation methods (e.g., block-based motion estimation and pel-recursive methods) that lie primarily in the former class [3]. Brown [4] and Zitova et al. [5] provide extensive surveys of image registration techniques covering methods in both classes. In this paper, we propose a new computationally efficient technique for image registration in the first class, based on adaptive filtering.

Adaptive filters have been successfully applied to a number of system-identification problems in the 1-D domain, a particular example being echo-cancellation [6]. In these applications, the adaptive filters not only allow the estimation of an unknown system but also incorporate the capability to track smoothly varying changes in the system. In this paper, we formulate image-registration as a 2-D system identification problem with spatially varying system parameters. Using the formulation, we motivate the development of a new image registration technique based on adaptive filtering. Since the successive update procedure in adaptive filtering is inherently 1-D, we map the 2-D image plane into a one-dimensional sequence using space-filling curves. This ensures spatial contiguity in the 2-D image plane, which is a pre-requisite for filter convergence and tracking.

The proposed adaptive filtering technique provides a method for local image registration that is capable of handling smoothly varying changes in registration between the input images. The method is computationally simpler than other methods for local image registration such as the pyramid-based image registration techniques. An additional benefit of the method is its reliance on only local information in each of the images. Both these features make the method well-suited for use in imaging sensor networks, where registration may be needed for mosaicking or super-resolution and memory and computational resources are scarce.

2. IMAGE REGISTRATION AS A SYSTEM IDENTIFICATION PROBLEM

Consider a pair of images $I_1(x, y)$ and $I_2(x, y)$ with overlapping views of the same scene, but with differences in the underlying geometry. Such images could be obtained, for instance, as successive frames of a video, multiple camera views, or multiple exposures from a single camera with camera displacement between exposures. Over the region of overlap, the pixel values in one image can be expressed in terms of the pixel values in the other image. In general terms, this relation may be expressed as a spatially varying system $h_o(x, y; x_o, y_o)$ which maps the geometry of image I_1 to I_2:

$$I_2(x_o, y_o) = \sum_{x,y} h_o(x, y; x_o, y_o) I_1(x, y) + e(x_o, y_o) \qquad (1)$$

For a number of imaging scenarios, this equation may be explicitly motivated by optical-flow models or through the use of camera models under suitable 3-D scene assumptions. In other scenarios, such as in the presence of camera distortions, physical arguments would justify the use of the above model.

The problem of image registration can now be regarded as a system identification problem, where the system response $h_o(\cdot)$ is

to be determined. We use this formulation and propose the use of adaptive filters for the estimation of $h_o(x, y; x_o, y_o)$.

2.1. 1-D Adaptive Filtering

For 1-D temporal signals, adaptive filtering is typically a two step process as shown in Fig. 1: i) a filtering process, 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 ii) 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) [6, Chap. 5] adaptation algorithm, the adaptive filtering process is given as:

$$\hat{v}_2(t_o) = \sum_{(t \in U)} \hat{h}(t, t_o) v_1(t) \tag{2}$$

$$e(t_o) = v_2(t_o) - \hat{v}_2(t_o) \tag{3}$$

$$\hat{h}(t + 1, t_o) = \hat{h}(t, t_o) + \beta e(t_o) v_1(t), \forall t \in U \tag{4}$$

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 changes in $h(t, t_o)$. The adaptation step-size β determines the speed of convergence, tracking capability, and the closeness of the approximation [6, Chap. 5-6].
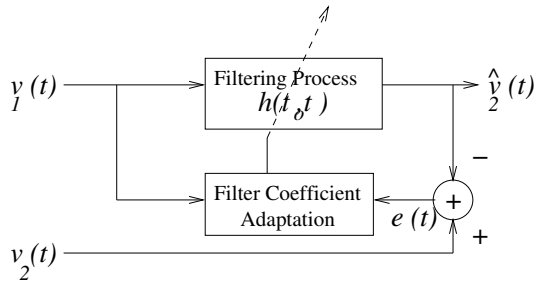

Fig. 1. Least-mean-square adaptation

2.2. 2-D Adaptive Filtering for Image Registration

For two-dimensional images, the adaptive filter takes the format of a 2-D finite impulse response (FIR) filter, h(x,y). Figure 2 shows the support of the filter, R, on the reference image, $I_1(x, y)$. Using the 2-D notations, the 2-D LMS adaptation algorithm can be written as:

1)Filter output (Prediction phase):

$$\hat{I}_2(x_o, y_o) = \sum_{(x, y \in R)} \hat{h}_b(x, y; x_o, y_o) I_1(x, y) \tag{5}$$

2)Estimation error:

$$e(x_o, y_o) = I_2(x_o, y_o) - \hat{I}_2(x_o, y_o) \tag{6}$$

3)Filter adaptation (Update phase):

$$\hat{h}_a(x, y; x_o, y_o) = \hat{h}_b(x, y; x_o, y_o) +$$
$$\beta e(x_o, y_o) I_1(x, y) \tag{7}$$

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

$$\hat{h}_b(x, y; x_n, y_n) = \hat{h}_a(x, y; x_o, y_o) \tag{8}$$

where $(x, y) \in R$ and β is the adaptation step-size. The subscripts 'b' and 'a' denote 'before' and 'after' adaptation, respectively. Under the appropriate conditions, (i.e., the right step-size and the right filter size) $\hat{h}(x, y; x_o, y_o)$ converges to the system model, $h_o(x, y; x_o, y_o)$ which maps the reference image to the current image.
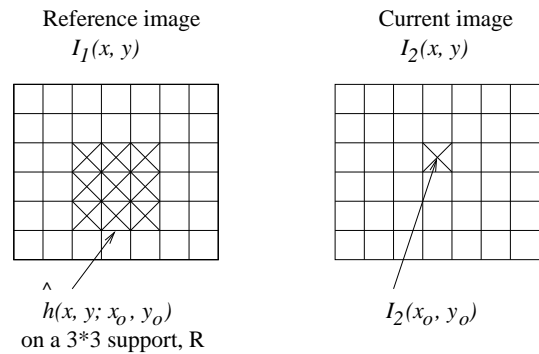

**Fig. 2**. 2-D adaptive filtering for images

2.3. Contiguity Preservation Using Space-filling Curves

It is important to exercise caution in selecting the scan order, i.e., the sequence in which the image pixels are visited, for the adaptive filtering process outlined above. Unlike 1-D signals, where time provides the natural ordering, there is no inherent ordering for the 2-D spatial case imposed by causality. However, 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. 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.

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 multi-dimensional space into the one-dimensional space. Because of its contiguity-preserving properties, Hilbert Curves [7] are used to determine the scanning order of pixels. Figure 3 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.

3. IMPLEMENTATION

In addition to the space-filling curves mentioned in the previous section, the implementation of the algorithm requires a number of
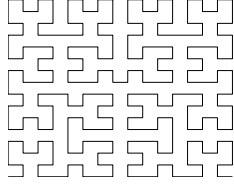
Fig. 3. Hilbert Curve on a 16×16 square

additional choices. To prevent the gradient noise amplification and increase the convergence rate, the normalized LMS adaptation algorithm [6, Chap. 6] is used. This results in a 2-D adaptive filter closer to the system model, and therefore better image registration. We also observe that unlike backwardly adaptive prediction methods used in compression, we have access to information in both images and can also utilize the current pixel information in the update process. However, if indiscriminate changes in filter coefficients are allowed, the use of the current pixel information results in an ill-posed problem, which needs regularization. This is achieved through a uniform motion constraint over a block of pixels around the current pixel, for which the adaptive filter is estimated.

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 cause a deterioration in convergence behavior in addition to the increased computational burden in adaptation. We therefore modify our algorithm to keep track of larger shifts independently as integer pixel shifts in x and y. This is achieved by constantly adjusting these values to ensure that the center of mass for the adaptive filter coefficients is located in the central region of our filter support.

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 the gradient estimation noise [6, Chap. 5].

Because the proposed registration technique is designed to solve local mis-registration, there is an initial registration phase to handle large scale global mis-registration between the images. For this purpose, a hierarchical parametric image registration method [8] is utilized to estimate the coarse registration among the input images.

4. EXPERIMENTAL RESULTS

We apply the proposed 2-D LMS adaptive filtering technique to register real-world image pairs. A digital camera is used to take pictures from an outdoor scene. Figures 4 and 5 show two such input images. Note that the captured scene in the images contains multiple objects located in different planes and thus the "planarity assumption" is invalid. In the absence of camera and scene depth orientation, the images cannot be readily registered using global

parametric models: an effect we describe as "locally varying motion". Additionally, the images exhibit differences in lighting due to deviations from lambertian assumptions.

The initial image registration among these two input images is performed using a hierarchical 8-parametric registration algorithm [8]. Figure 6 shows the absolute registration error image after the initial registration. The estimated motion parameters are used as the initial adaptive filter, for the first pixel in the scan order of the current image. Then, the adaptive filter is updated following the scan-order established by the Hilbert Curve, using the proposed 2-D adaptive filtering technique.

Since the Hilbert Curve requires the 2-D region to be a square, the input images are divided into overlapping squares. Then, a Hilbert Curve is computed for each square. The proposed 2-D adaptive filtering technique is performed on each square iteratively. The estimated filter coefficients on one square are used as the initial adaptive filter on the next square in such a way that the scanning order of the whole input image is guaranteed to be contiguity-preserving.

Throughout the input image, a constant size (13×13) adaptive filter with an adaptation step-size of 0.2 is used. Figure 7 shows the converged filter at one location. In order to prevent ill-posedness, a uniform motion constraint is put around each pixel over a block of 3*3. Figure 8 shows the absolute registration error image after the proposed 2-D LMS adaptive filtering technique is applied. Both Figures 6 and 8 are contrast-enhanced (i.e., multiplied by 4) for better presentation. The pSNR values of the error images for the parametric registration algorithm and the proposed technique are 19.98 dB and 29.45 dB, respectively. The proposed image registration technique is also compared with the pyramid-based Lucas-Kanade technique [1]. The corresponding pSNR of the absolute registration error image for the Lucas-Kanade technique is computed to be 25.73 dB.

Fig. 4. Reference input image

5. CONCLUSION

In this paper, we propose a new signal-processing framework for image registration. Formulating the problem as the 2-D analog of the system-identification problem, we develop a (spatial) adaptive
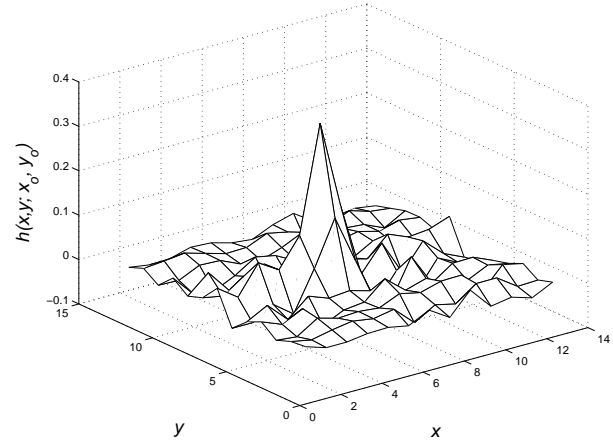
Fig. 5. Current input image

Fig. 7. 2-D adaptive filter

Fig. 6. Registration error after parametric model registration

Fig. 8. Registration error after the proposed registration method

filtering algorithm for image registration. The method operates by using the adaptive filter to predict one image from the other, one pixel at a time, and updating the filter coefficients after each prediction. Once the filter converges to the underlying system response, the process also establishes correspondences among the images. By using a suitable space-filling curve for the traversal of pixel-locations in the image, we ensure that the system to be tracked by the filter is slowly varying, thus achieving convergence and improved tracking/noise performance. The method has the benefit of requiring only local information at each point, which makes distributed implementation feasible, and also has a lower computational requirement than other methods for locally varying registration of images.

6. REFERENCES

[1] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. DARPA Image Understanding Workshop*, 1981.

[2] David A. Forsyth and Jean Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, Englewood Cliffs, NJ, 2002.

[3] A. M. Tekalp, Ed., *Digital Video Processing*, Prentice Hall, Upper Saddle River, NJ, 1995.

[4] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

[5] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[6] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, NJ, 2002.

[7] H. Sagan, *Space-filling curves*, Springer, Berlin, 1994.

[8] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," *European Conference on Computer Vision*, pp. 237–252, May 1992.