

ROBUST IMAGE REGISTRATION USING LOG-POLAR TRANSFORM

George Wolberg

Siavash Zokai

Department of Computer Science

City College of New York

New York, NY 10031

{wolberg|zokai}@cs-mail.engr.cuny.cuny.edu

ABSTRACT

This paper describes a hierarchical image registration algorithm for affine motion recovery. The algorithm estimates the affine transformation parameters necessary to register any two digital images misaligned due to rotation, scale, shear, and translation. The parameters 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. A log-polar registration module is introduced to accommodate arbitrary rotation angles and a wide range of scale changes. This serves to furnish a good initial estimate for the optimization-based affine registration stage. We demonstrate the hybrid algorithm on pairs of digital images subjected to large affine motion.

1. INTRODUCTION

Image registration refers to 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. 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. This is of practical importance in many fields, including remote sensing, medical imaging, and computer vision [1].

This paper describes a hierarchical image registration system based on parameter estimation techniques. We model the mapping function as an affine transformation. The algorithm estimates the affine parameters necessary to register any two digital images misaligned due to rotation, scale, shear, and translation. The parameters are selected to minimize the sum of squared differences 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 a wide range of scale changes. Its purpose is to furnish a good initial estimate to the affine registration module that is based on nonlinear least squares optimization.

This hybrid approach yields superior results over either module acting independently. We demonstrate the algorithm by recovering the affine parameters for pairs of digital images subjected to large affine motion.

2. PREVIOUS WORK

This paper addresses image registration under affine transformations. Recent work in this area has focused on solutions in the frequency domain. Registration problems involving pure translation can be recovered by applying the Fourier transform and using phase correlation [2]. A Fourier-based method to accommodate translation and rotation was described in [3].

The Fourier-Mellin transform has been introduced to register images that are misaligned due to translation, rotation, and scale [4, 5, 6, 7]. This method applies a Fourier transform to images to recover translation. Then a log-polar transformation is applied to the magnitude spectrum and the rotation and scale is recovered by using phase correlation in the log-polar space. This method exploits the fact that by operating on the magnitude spectrum of an image, the translational differences are avoided since the magnitude spectrum of an image and its translated counterpart are identical; only their phase spectrum are different. Furthermore, the log-polar transformation causes rotation and scale to be manifest as translation, whereby phase correlation can be applied to recover the rotation angle and scale factor between the pair of input images. The problem here, though, is that

Appears in *Proc. of IEEE Intl. Conf. on Image Processing*, Sep. 2000.
This work was supported by a NASA FAR Award (NAG-57129).

limited scale factors can be determined because large scale factors would alter the frequency content beyond recognition. It should be noted that the maximum scale factor recovered in [6] and [7] is 2.0 and 1.8, respectively.

The work presented in this paper consists of two modules: log-polar registration followed by optimization-based affine registration. The latter module is similar in spirit to the method described in [8]. However, since it is based on optimization techniques, it may fail to register an image pair unless they are marginally misaligned. We have addressed this problem by introducing a log-polar module to furnish a good initial estimate, even in the presence of arbitrary rotation angles and a wide range of scale changes. By operating in the spatial domain under a multiresolution framework, we have been able to recover a wider range of scale factors than those possible under the Fourier-Mellin transform.

3. AFFINE PARAMETER ESTIMATION

In this paper, we shall address the problem of registering images misaligned due to an affine transformation. The mapping function is given as

$$u = a_1x + a_2y + a_3 \quad (1)$$

$$v = a_4x + a_5y + a_6 \quad (2)$$

The six unknown parameters relating two input images, I_1 and I_2 , will be estimated by minimizing the sum of squared differences between I_1 and the transformed I_2 . The sum of squared differences (SSD) objective criterion establishes a similarity measure between two images:

$$\begin{aligned} \chi^2(\mathbf{a}) &= \sum_{i=1}^N [I_1(\mathbf{x}_i) - I_2'(\mathbf{x}_i)]^2 \\ &= \sum_{i=1}^N [I_1(\mathbf{x}_i) - T_{\mathbf{A}}\{I_2(\mathbf{u}_i)\}]^2 \\ &= \sum_{i=1}^N [I_1(\mathbf{x}_i, \mathbf{y}_i) - I_2(\mathbf{u}_i, \mathbf{v}_i)]^2 \end{aligned} \quad (3)$$

where $\mathbf{u}_i = a_1\mathbf{x}_i + a_2\mathbf{y}_i + a_3$ and $\mathbf{v}_i = a_4\mathbf{x}_i + a_5\mathbf{y}_i + a_6$, respectively. Note that T is a geometric transformation applied to image I_2 to map it from its $[\mathbf{u}, \mathbf{v}]$ coordinate system to the $[\mathbf{x}, \mathbf{y}]$ coordinate system of I_1 . The subscript \mathbf{A} denotes that the transformation is affine.

We adopt the Levenberg-Marquadt nonlinear least squares optimization algorithm [8, 9] to estimate the six unknown affine parameters.

4. LOG-POLAR REGISTRATION

In this section, we introduce a log-polar registration module that serves as a preprocess to the parameter estimation module. Although the parameter estimation method features sub-pixel 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 newly added module is to account for large geometric transformations, bringing images into close alignment even in the presence of large (ten-fold) scale changes, as well as arbitrary rotations and translations. In practice, we don't expect scale changes beyond factors of four, however arbitrary rotation angles and translations must surely be addressed.

We now review log-polar coordinate transformations and describe their use in registration. Consider the polar (r, α) coordinate system, where r denotes radial distance from the center (x_c, y_c) and α denotes angle. Any (x, y) point can be represented in polar coordinates:

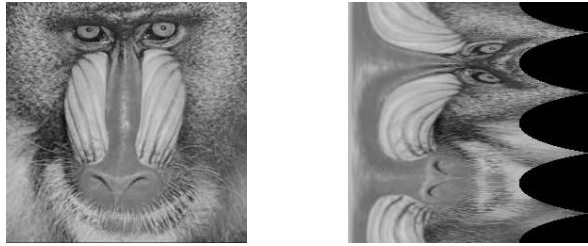
$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \quad (4)$$

$$\alpha = \tan^{-1} \left(\frac{y - y_c}{x - x_c} \right) \quad (5)$$

Applying a polar coordinate transformation to an image I maps radial lines in Cartesian space to horizontal lines in the polar coordinate space. We shall denote the transformed image I_p . If we assume that r and α lie along the horizontal and vertical axes, respectively, then image I shown in Fig. 1(a) will be mapped to image I_p in Fig. 1(b) after a polar coordinate transformation. Note that the origin in both coordinate systems is taken to be in the upper left corner. The benefit of this new coordinate space is that simple scale and rotation changes may be induced by modifying the (r, α) data. For instance, a circular shift along the α -axis in I_p induces a rotation of I . Recall, after all, that the data stored along the rows in $r\alpha$ -space represent radial data in xy . By moving the I_p rows up or down, that radial data will map to a new (rotated) set of radial lines. Fig. 1(c) shows a circular shift applied to Fig. 1(b). The resulting image in Cartesian space is shown in Fig. 1(d). Notice that since we shifted the 512-row image in Fig. 1(b) by 128 rows, the image in Fig. 1(a) has been rotated by 90° . Similarly, resampling in the r -axis can effect scale changes.

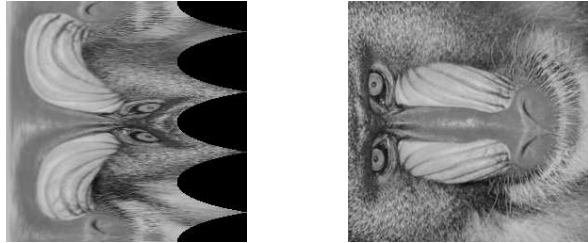
Consider two images I_1 and I_2 and their polar counterparts I_{1p} and I_{2p} . If I_2 is a rotated version of I_1 , then I_{2p} will be a circularly shifted version of I_{1p} along the α -axis. By applying cross-correlation to I_{1p} and I_{2p} , the offset $d\alpha$ can be found that best matches the two images. Recall that cross-correlation is commonly used for finding translational offsets (phase shifts) between two images. It does not work well in the presence of scale or rotation. However, in the polar coordinate space, finding the translational component $d\alpha$ between I_{1p} and I_{2p} corresponds to finding the rotation between I_1 and I_2 . There would be, for instance, no problem in computing the phase shift between Figs. 1(b) and 1(c).

A similar result is needed to determine the scale change between I_1 and I_2 . Consider a four-fold magnification of image I_1 , yielding an enlarged image I_2 . All points (x, y) in I_1 now map to $(4x, 4y)$ in I_2 . To determine the scale



(a) input image

(b) polar transformation



(c) circular shift in polar space (d) result in Cartesian space

Figure 1: Polar coordinate transformation.

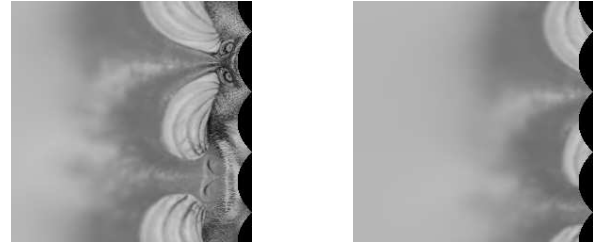
factor, we introduce the use of logarithms. In log-space, $(x, y) \rightarrow (\log x, \log y)$, and $(4x, 4y) \rightarrow (\log 4x, \log 4y) \rightarrow (\log x + \log 4, \log y + \log 4)$. It now becomes apparent that in log-space, the introduction of a scale factor manifests itself as a phase shift in the log-transformed image. Therefore, instead of mapping an image into (r, a) coordinate space, it becomes useful to map it into $(\log r, a)$ coordinate space using a log-polar transformation. Ordinary cross-correlation in the log-polar space now determines the best $d(\log r)$ and da phase shifts, which translates to scale and rotation in Cartesian space.

Fig. 2 shows an example. Figs. 2(a) and 2(b) depict two input images. The latter figure is a scaled, rotated, and cropped version of Fig. 2(a). The result is consistent with sensor movement that causes a four-fold zoom and a 45° rotation. Their log-polar transformations are shown in Figs. 2(c) and 2(d). Clearly, they differ by a phase shift. Compare, for instance, the left half of Fig. 2(c) with the right half of Fig. 2(d). When ordinary cross-correlation is applied to those two images, the computed phase shift accurately reflects the four-fold scale factor and 45° rotation that constitutes the transformation between the two input images.

We have used the log-polar transformation to recover differences as large as ten-fold scale changes. Given this superior performance over the state-of-the-art in frequency domain solutions, we believe that this approach should be exploited. There is one difficulty, though, that remains to be solved: finding the origin of the image from which the radial distance is measured. In the examples above, we have



(a) input image

(b) scale=4; rotation= 45° 

(c) log-polar transform of (a) (d) log-polar transform of (b)

Figure 2: Log-polar coordinate transformation.

assumed that the origin of both images lie at their geometric centers. In fact, their centers can be displaced and unless that correspondence (translation) is known, the information derived from the polar transformation is of limited value.

We have implemented a new algorithm for automatically finding the translation between both input images in the presence of scale and rotation. 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. The approach at any given level is outlined below.

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 :
 - Crop region I'_2
 - Compute I'_{2p}
 - Cross-correlate I'_{1p} and $I'_{2p} \rightarrow (dx, dy)$
 - If maximum correlation, save (x, y) and (dx, dy)
4. Scale $\leftarrow dx$
5. Rotation $\leftarrow dy$
6. Translation $\leftarrow (x, y)$

5. RESULTS

Fig. 3 shows an example of the use of log-polar registration as a preprocess for the sub-pixel perspective registration module. Figs. 3(a) and 3(b) show two images of a building taken with different rotation and focal length parameters. Their log-polar transformations are shown in Figs. 3(c) and 3(d). Cross correlation in the log-polar domain produces a phase shift that corresponds to a scale factor of 0.469 and a rotation of -201° . The translation is (-11,-51). Fig. 3(e) shows the overlay results of log-polar registration after Fig. 3(b) was scaled, rotated, and translated to best match Fig. 3(a). This result is used as an initial estimate for the subsequent affine parameter estimation module. The affine transformation derived *after* log-polar registration is given below.

$$u = 1.008x + 0.006y - 1.669 \quad (6)$$

$$v = 0.003x + 1.045y + 0.203 \quad (7)$$

Note that log-polar registration was able to bring the two images into sufficient alignment that the subsequent affine transformation is nearly an identity mapping. The resulting overlay is presented in Fig. 3(f). Notice the improved registration results as evident by the reduced double exposure effect in the overlaid images.

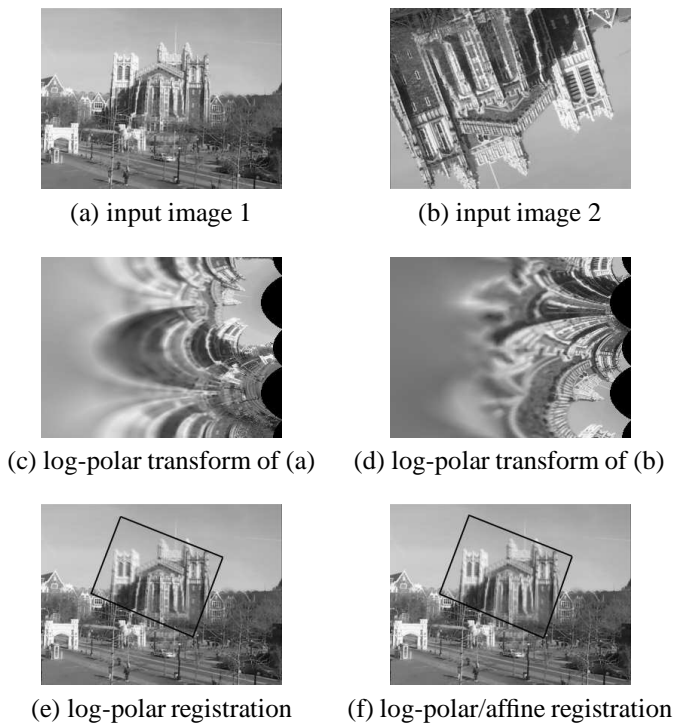


Figure 3: Polar coordinate transformation.

6. CONCLUSIONS

This paper has presented a hierarchical image registration algorithm to register any two digital images misaligned due to rotation, scale, shear, and translation. The algorithm couples the log-polar transform with a nonlinear least squares algorithm to estimate the affine transformation parameters. Although the Fourier-Mellin transform also uses the log-polar transformation to recover rotation and scale, it is limited in use to two-fold scale factors. Larger scale changes induce too much distortion to the Fourier coefficients to be useful for affine recovery. Instead, this work operates directly in the spatial domain and 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 affine registration module based on nonlinear least squares optimization. That module, based on the work in [8], offers subpixel precision. Coupling the two modules in this manner facilitates the registration of images in the presence of large-scale affine transformations. Future work will investigate extensions of this approach for the recovery of perspective parameters.

7. REFERENCES

- [1] Lisa G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, December 1992.
- [2] C.D. Kuglin and D.C. Hines. The phase correlation image alignment method. *Proc. Int. Conf. on Cybernetics and Society*, pages 163–165, 1975.
- [3] E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, (3):700–703, September 1987.
- [4] D. Casasent and D. Psaltis. Position, rotation, and scale-invariant optical correlation. *Applied Optics*, 15:1793–1799, 1976.
- [5] W.K. Pratt. *Digital Image Processing*. John Wiley & Sons, New York, 1978.
- [6] Q. Chen, M. Defrise, and F. Deconinck. Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16(12):1156–1168, December 1994.
- [7] B.S. Reddy and B.N. Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 5(8):1266–1270, August 1996.
- [8] Philippe Thévenaz, Urs E. Ruttimann, and Michael Unser. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Processing*, 7(1):27–41, January 1998.
- [9] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 2nd edition, 1993.