Morphological Pyramid Image Registration

Zhongxiu Hu and Scott T. Acton
The Oklahoma Imaging Laboratory
School of Electrical and Computer Engineering
Oklahoma State University
Stillwater, OK 74078
Email: hzhongx@fajita.ecen.okstate.edu, sacton@ceat.okstate.edu

Abstract

We proposed an intensity-based morphological pyramid image registration algorithm. This approach utilizes the global affine transformation model, also considering radiometric changes between images. With the morphological pyramid structure, Levenberg-Marquardt optimization, and bilinear interpolation, this algorithm can be implemented hierarchically and iteratively with capability of measuring, to subpixel accuracy, the displacement between images subjected to simultaneous translation, rotation, scaling, and shearing. The morphological pyramid shows better performance than Gaussian pyramid in this matching technique.

1. Introduction

Image registration is the process of matching two or more images of the same scene taken at different times, from different viewpoints, or by different sensors. It is a fundamental task in automated video tracking, remote sensing analyses, and sensor fusion. The goal of image registration is to establish the spatial correspondence between two images. Several techniques have been developed for various types of data and applications [1]. Existing techniques fall into two categories: intensity-based approaches [2, 3] and feature-based techniques [4, 5]. In order to obtain a highly efficient and robust algorithm, pyramidal architectures are commonly utilized in the registration process [2, 3, 5, 6].

In this paper, we present a *Morphological Pyramid Image Registration* (MPIR) algorithm that uses an intensity based differential method for matching. This algorithm considers a model combining a 2-D affine transformation and an illumination change. The multi-resolution images are represented by a *Morphological Pyramid* (MP), as the MP's have the capability to eliminate details and to maintain shape features. The *Levenberg-Marquardt* nonlinear optimization algorithm is employed to estimate the matching parameters. A bilinear interpolation is performed to resample images and their derivatives required for the transformation. This algorithm is capable of measuring, to subpixel accuracy, the displacement between images subjected to simultaneous translation, rotation, scaling, and shearing. The benefits of this method are the accuracy and stability of estimation, the automated solution, and the low computational cost without pre-estimation. The novelty of this method lies in the use of MP's for image registration, which improves the performance over the *Gaussian Pyramid*.

This work was supported by NASA EPSCoR.

2. MPIR Model

The complete model for image matching techniques should consider the geometric relation and intensity relation between images [7]. These correspondences reflecting the total relationship between images can be described by mapping functions with unknown parameters. Image registration aims at estimating and evaluating the parameters within the considered model.

2.1. Global Affine Transformation

The spatial-mapping function and parameters in MPIR are described by a global affine transformation, which is commonly used in image registration, since the image projections from different viewpoints can be well approximated by using this geometric transformation. The general affine transformation includes translation (tx, ty), rotation (θ), scaling (sx, sy), and shearing (shx, shy) [1] and can be expressed as

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ shy & 1 \end{pmatrix} \begin{pmatrix} 1 & shx \\ 0 & 1 \end{pmatrix} \begin{pmatrix} sx & 0 \\ 0 & sy \end{pmatrix} \times \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} r \\ c \end{pmatrix} + \begin{pmatrix} tx \\ ty \end{pmatrix}. \quad (1)$$

With six parameters, we can simplify the above equation:

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_4 & a_5 \end{pmatrix} \begin{pmatrix} r \\ c \end{pmatrix} + \begin{pmatrix} a_3 \\ a_6 \end{pmatrix}. \qquad (2)$$

The affine transformation can accommodate shearing in addition to scaling, translation, and rotation.

2.2. Radiometric Changes

In most practical cases, consideration of illumination changes is sufficient, but it may also be necessary to compensate for brightness and contrast between images, which are caused by the radiometric variation in imaging. The intensity-mapping function and parameters in MPIR take care of the changes in brightness and contrast and are expressed by the following model:

$$g_2 = a_7 g_1 + a_8 . \qquad (3)$$

2.3. Intensity-based Differential Model

Combining the spatial mapping and the intensity mapping functions, we achieve the complete relationship between the two input images:

$$g_2(r,c) = [a_7 g_1(p,q) + a_8] + n(r,c) , \qquad (4)$$

where $n(r,c)$ is due to noise existing in both images, and the eight transformation parameters a_k , k = 1, 2, … , 8, are unknown. The MPIR algorithm utilizes the intensity-based method for matching, since registration methods based on initial intensity values can make effective use of all data available. The objective of MPIR is to estimate these eight unknown parameters.

Given the approximate initial values a_{0k} for each a_k , which give the corresponding point (p_0, q_0) by (2), we linearize g_2 around the point (p_0, q_0) and only consider the differential linear model resulting in the following equations:

$$g_2(r,c) \cong \hat{g}_1(p_0, q_0) + J^T \Delta A + n(r,c) \qquad (5)$$

with $\hat{g}_1(p_0, q_0) = a_{07} g_1(p_0, q_0) + a_{08}$,

$$\Delta A = A - A_0 = [\Delta a_k] \qquad (6)$$

$$= [a_1 \ \ . \ \ . \ \ . \ \ a_8]^T - [a_{01} \ \ . \ \ . \ \ . \ \ a_{08}]^T ,$$

$$f_r = \frac{\partial \hat{g}_1(r,c)}{\partial r} |_{(p_0,q_0)} , \qquad f_c = \frac{\partial \hat{g}_1(r,c)}{\partial c} |_{(p_0,q_0)} ,$$

$$f = g_1(r,c) |_{(p_0,q_0)} ,$$

and $J = [f_r r \quad f_r c \quad f_r \quad f_c r \quad f_c c \quad f_c \quad f \quad 1]^T$.

From (5), the sum of squared errors (SSE) between two images can be obtained as

$$\Omega = \sum n^2(r,c) = \sum \left(g_2 - \hat{g}_1 - J^T \cdot \Delta A\right)^2$$

$$= \sum (\Delta I - J^T \cdot \Delta A)^2 . \qquad (7)$$

We then minimize the SSE with respect to the eight unknown parameters, Δa_k (k = 1, 2, …, 8), and, thus, estimate the parameters by

$$\Delta A = \left[\sum J \cdot J^T\right]^{-1} \cdot \sum \Delta I \cdot J , \qquad (8)$$

which yields the six corrections Δa_k , k =1, 2, … , 6 to the approximate values of geometric transformation and the corrections Δa_k , k =7, 8 for the radiometric parameters. Eq. (8) reveals that the estimation of parameters is independent of the noise and can be estimated simply by the difference between the two images at the corresponding positions and the partial derivatives of image g_1 . The edge pixels contribute more than the other pixels to the estimation process. Finally, we can update the transformation parameters by using (6) with the estimated increment ΔA .

3. The MPIR Algorithm

The pyramid architecture offers a framework for image registration with decreased computational cost and increased solution quality. The image pyramid is a multi-resolution representation of an image constructed by successive filtering and sub-sampling. It allows scale selection appropriate resolution for the task at hand. For registration, the differential linear model discussed here is only valid for subpixel displacements, but employing a pyramid structure can increase the functionality to displacements of several pixels. Moreover, the use of a multi-resolution scheme avoids some local optima in the solution (because they do not exist at coarse pyramid representations) and often widens the capture area of the global optimum. From discussion above, we know that edge pixels take greater roles than others do in the differential method. As the inherent goals of morphological operations are the preservation of region shapes and elimination of irrelevancies, the morphological pyramid is a reasonable choice for the differential matching method.

3.1. Morphological Pyramids

Mathematical morphology is a set-theoretic approach to image analysis. It provides a quantitative description of the geometric structures of an image. The morphological filters, such as open and close, are relatively inexpensive nonlinear filters. The filters can designed to preserve edges or shapes of objects, while eliminating noise and details in an image.

In generating the MP, the morphological open-close and close-open filters are typically chosen because they are biased-reduced operators [8]. The MP of an image can be constructed by successively morphological filtering and subsampling:

$$I_L = \lfloor (I_{(L-1)} \circ K) \bullet K \rfloor_{\downarrow d} \quad L = 0,1, \ldots, n, \qquad (9)$$

where L is called the pyramid level. I_0 is the original image, $[\cdot]_{\downarrow d}$ represents a down sampling by a factor of *d* in each spatial dimension (along rows and columns), $(I \circ K)$ represents the morphological opening of the image *I* with structuring element *K*, and $(I \bullet K)$ represents the morphological closing. The finest level $L = 0$ of the *MP* contains the input image. The image at any level L is created by applying the morphological close-open filter with a 3x3-element structure to the image at level $(L-1)$ and then subsampling the filtered image with $d = 2$.

3.2. The Levenberg-Marquardt Algorithm

After we compute the image pyramids, we use Levenberg-Marquardt (LM) algorithm to estimate the transformation parameters iteratively. The LM nonlinear optimization algorithm is well suited for performing registration based on a least-squares criterion [9]. The following equation gives the estimate corrections

$$\Delta A_K = \left[(\sum J_K \cdot J_K{}^T) + \mu_K U_K \right]^{-1} \cdot \sum \Delta I_K \cdot J_K \quad (10)$$

for K^{th} step, where μ_K is a changeable parameter and U_K is an identical matrix (8 x 8). The LM algorithm provides a compromise between the speed of Newton's method and the guaranteed convergence of steepest descent.

3.3. Bilinear Interpolation

For each step of estimate process, we must resample one image to update the transformed version of the registered image and its gradients. The simplest scheme for gray-level interpolation is based on a nearest neighbor approach called zero-order interpolation. But the nearest neighbor interpolation yields undesirable artifacts such as the distortion of straight edges. Smoother results can be obtained by using more sophisticated techniques, such as bicubic interpolation. However, this technique is computationally expensive. Bilinear interpolation is a reasonable compromise between smoothness and computational cost. With this method, the intensity at (p,q) can be computed by using the gray levels of the four nearest neighbors, *a*, *b*, *c*, and *d*, in

$$g(p,q) = a \cdot p + b \cdot q + c \cdot p \cdot q + d \,. \qquad (11)$$

3.4. Matching Measures

The MPIR technique is based on minimizing the SSE between two images. Therefore, the matching measures are based on measuring the SSE. In practice, we consider three matching criteria for the stopping criteria. The first criterion is an absolute error $E = \sum \left| g_2(r,c) - \hat{g}_1(p_0, q_0) \right|$ less than the threshold T_1. The second measure uses the observed relative gain $\Delta E / E$ at each successive iteration step. When this gain is below a priori threshold T_2, the convergence is reached. The third criterion compares the maximum corrections of the transformation parameters $\max\{ |\Delta a_k|, k = 1,2,\ldots,8 \}$ to a threshold T_3.

3.5. Matching Procedure

We introduce an efficient optimization scheme that uses a coarse-to-fine iterative refinement strategy over the MP pair. Within each pyramid level, three matching criteria are used to indicate convergence of the matching process. Once convergence has been achieved at a particular pyramid level, a transition to a finer level is made. The solution from the previous pyramid level is used as an initial estimate. The new incremental estimate is computed by minimizing the SSE between the resampled images $\hat{g}_1$ and image g_2. The estimated increment is then composed with the previous estimate to achieve a new estimate. This whole process is iterated at each pyramid level and then across pyramid levels to achieve the final estimation. To verify the matching robustness and accuracy, we implement both Forward Matching (FM) and Backward Matching (BM) for each image pair. In FM, we register the first image with the second image, but reverse the order in BM (*viz.*, the second image is registered with the first image).

4. Experimental Results

The proposed algorithm has been tested on artificial image pairs and natural image pairs. The initial estimated parameters are set to

$$A = [1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0]^T$$

for all matching processes.

Fig. 1 shows an example with the image "Lena." Fig. 1(b) is obtained by transforming Fig. 1(a) with Eq. (1)-(3) and the following parameters: translation (*tx*=5.09, *ty*=-3.01), rotation (θ=-7.70), scaling (*sx*=0.91, *sy*=1.25), shearing (*shx*=0.05, *shy*=-0.15), contrast (a_7=1.2), brightness (a_8=4.05). These data provide the actual transformation parameters represented as "Act." in Table 1. Fig. 1(c) shows the FM result of registering Fig. 1(a) to Fig. 1(b).

The matching results with the MPIR algorithm for "Lena" are summarized in Table 1, where the row AFE (or ABE) stands for the absolute error between the actual parameters and FM (or BM) results. FBE is the error between the FM and BM estimated parameters. For example, with FM and BM parameters p_F and p_B, we have $FBE = \left| p_F - p_B \right|$. The last two columns in the table show the maximum deformation errors along x^{th} row and y^{th} column. All error values show us that the matching algorithm works very well with subpixel accuracy. Although there are significant errors for the radiometric parameters, the errors for 256 gray-levels are relatively small.

Another example of MPIR is presented in Fig. 2 using the real image pair "Old Central." The original images are displayed in Fig. 2 (a)-(b) followed by the FM version of Fig. 2(a) shown in Fig. 2(c). The matching results with small errors are given in Table 2. In order to show that the morphological pyramid matching results are more accurate than that of single resolution, we list the matching errors for single resolution with the last row (SE) in Table 2.

To further verify the performance of the MPIR algorithm, we also implement both FM and BM for each image pair with the Gaussian Pyramid (GP). Table 3 presents the errors for image pairs "Lena" and "Old Central," where MAE and GAE indicate the errors between the actual parameters and the estimated parameters with the MP and GP, respectively. Here, ME and GE represent the errors between the FM and BM parameters with respect to the MP and GP. From the last two columns in the table, we can see that both of the pyramid structures work well for the image pair "Lena." However, the GP fails for "Old Central." The errors are caused by the edge shifting due to Gaussian filtering. The registration results with GP are also shown in Fig. 3(b)-(c), where the FM result in Fig. 3(c) appears to be successful, but the BM result in Fig. 3(b) is not indicative of success. Comparing the matching results between the MP and GP, we can infer that MP gives improvements in accuracy and robustness.

5. Conclusion

The MPIR algorithm with an intensity-based differential matching technique is reliable and efficient. The MP architecture improves robustness while decreasing the likelihood of being trapped at a false local optimum. MPIR increases the matching range of the differential method, and reduces the computational cost. With LM nonlinear optimization and bilinear interpolation, this algorithm is capable of measuring, to subpixel accuracy, the displacement between images subjected to affine transformation, which includes simultaneous translation, rotation, scaling, and shearing. The matching procedure is entirely automatic and does not utilize pre-estimation.

6. References

[1] L.G. Brown, "A survey of image registration techniques," *ACM Computer Surveys*, vol. 24, pp. 325-376, 1992.

[2] P. Anandan, "Model based techniques for image registration and three dimensional scene analysis from image sequences," in *Proc. of the 27th Asilomar IEEE Conf. on Signals, Systems and Computers*, pp. 90-94, 1993.

[3] P. Thevenaz, U.E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Processing*, vol.7, pp. 27-41, 1998.

[4] H. Li, B. S. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. Image Processing*, vol.4, pp. 320-334, 1995.

[5] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Trans. Image Processing*, vol.2, pp. 311-326, 1993.

[6] V. Cantoine and S. Levialdi, *Pyramid Architectures for Computer Vision*, Plenum Press, New York, 1994.

[7] R. Haralick, and L. Shapiro, *Computer and Robot Vision, Volume 2*. Addison-Wesley, 1993.

[8] A. Morales, T. Acharya, and S. Ko, "Morphological pyramids with alternating sequential filters", *IEEE Trans. Image Processing*, vol.4, pp. 965-977, 1995.

[9] M. Hagan, H. Demuth, and M. Beale, *Neural Network Design*, PWS Publishing Company, 1996.

Table 1. MP matching results for "Lena"

I	ΔA	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	Max. Δx (pix.)	Max. Δy (pix.)
L	**Act.**	0.8934	0.1839	5.0900	-0.3015	1.2111	-3.0100	1.2000	4.0500		
E	***FM***	0.8934	0.1839	5.0900	-0.3015	1.2111	-3.0099	1.2000	4.0504		
N	***BM***	0.8934	0.1838	5.0943	-0.3015	1.2111	-3.0081	1.1760	6.4270		
A	***AFE***	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0004	0.0000	0.0001
	ABE	0.0000	0.0001	0.0043	0.0000	0.0000	0.0019	0.0240	2.3770	0.0172	0.0089
	FBE	0.0000	0.0001	0.0043	0.0000	0.0000	0.0018	0.0240	2.3766	0.0172	0.0089

Table 2. MP matching results for "Old Central"

I	A	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	Max. Δx (pix.)	Max. Δy (pix.)
O	***FM***	0.9781	-0.2150	-10.6672	0.2131	0.9789	-6.9803	1.0097	-4.1355		
C	***FB***	0.9781	-0.2150	-10.6704	0.2131	0.9788	-6.9838	1.0210	-5.8815		
	FBE	0.0000	0.0000	0.0032	0.0000	0.0001	0.0035	0.0113	1.7460	0.0133	0.0000
	SE	0.0015	0.0014	0.0630	0.0002	0.0009	0.0580	0.0171	2.5789	0.4324	0.0000

Table 3. Comparison between MP and GP matching results

I	*A*	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	Δx	Δy
L	*MAE*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0004	0.0000	0.0001
	GAE	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000	0.0011	0.0003	0.0007
O	*ME*	0.0000	0.0000	0.0032	0.0000	0.0001	0.0035	0.0113	1.7460	0.0133	0.0000
	GE	0.1248	0.1350	2.8422	0.1549	0.0207	4.6797	0.1859	29.8201	34.091	0.0000

Fig. 1: Registration of image pair "Lena": (a) Original Lena, (b) Transformed, (c) Registration of (a) to (b)

Fig. 2: Registration of image pair "Old Central": (a) Old Central 1, (b) Old Central 2, (c) Registration of (a) to (b)

Fig. 3: Comparison of MP and GP matching results: (a) BM result (MP), (b) BM result (GP), (c) FM result (GP)