AN IMAGE REGISTRATION ALGORITHM BASED ON DIGITAL IMAGE FUSION

WANG Rong, GAO Li-qun

Key Laboratory of Process Industry Automation, Ministry of Education, China, Shenyang 110004,China

Abstract: In this paper, an image registration algorithm based on both feature and intensity is proposed. Decomposing the source images with wavelet, corner points can be detected in the two approximate sub-images. The initial registration parameters can be achieved through affine transformation model, the coarse registration parameters can be obtained by the optical flow estimating and iterative refining, and the parameters are passed on to the next finer level to refine. The process will be repeated until we obtain the final registration parameters. Theory and experiments show this approach provided accurate registrations for DCIF images.

Keywords: Image fusion; Image registration; Corner point detection; Multi-scale

1. INTRODUCTION

The image registration technique is applied in many fields, such as medical images, object recognition, vision of robot, processing of remote sensing images, image fusion. Image registration is the key problem in image fusion technique.

In the recent 20 years, researchers had put forward many different image registration methods used for various types of sensors and application fields. The previous registration techniques are the methods based on correlation, the shortage of which is the complex computation (Pratt, 1974). Recently, registration methods based on features and areas were put forward by some researchers (Goshtasby, 1986). Existing registration methods can be classified into two categories: the intensity-based method and the feature-based method (Rogelj, 2003). In the former, the images were registered by selecting a number of windows in high-variance intensity areas of one image, by locating the corresponding windows in the other image and by taking the window's geometric center or mass center as control points to determine the registration parameters. In the latter, the features (edge, corner point, contour) were extracted from source images and registered according to the common features, and the registration parameters can be determined. In our study, we found that neither of the above image registration approaches is well suited to a digital camera image fusion (DCIF). For example, if the two images were obtained using different focus, a feature which appeared stronger in one image may appear weaker in other image. The high-variance area in one image will not always correspond to the one in other image. This leads to difficulties in finding control points for the intensity-based method. Similarly, a common feature may appear to look quite different in the two images, which makes feature matching nearly impossible. In view of above-mentioned, in this paper we proposed a hierarchical scheme, which uses both feature-based and intensity-based methods. Meanwhile, we also combine the idea of optical flow with a coarse-to-fine multi-resolution method to overcome some of the limitations of the intensity-based scheme. Experiments show that this algorithm provided accurate registrations for the DCIF application.

2. FEATURE-BASED INITIALIZATION

In this paper, we proposed a way to combine the optical flow estimation with the feature-based registration which provides good effect for DCIF application. First, a multi-scale wavelet decomposition is applied to the two source images to

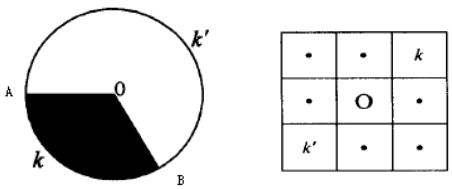
obtain two sets of wavelet pyramid. A corner point detection is performed on low frequency sub-images at the coarsest level of the two source images. The corner point pair adopted is used as the major image features for the initial matching to get the initial registration parameters. After several iterative refinement steps of the optical flow estimation, the best registration parameters of this level are achieved. Then take this parameters as the initial registration parameters of the next level and repeat the same iterative refinement several times until the finest decomposition level is reached and the final accurate registration parameters are obtained.

2.1 Corner point detection

The aim of corner point extraction is to detect the corner point, edge and the edge's orientation and inclusive angle in the images. Scanning along the arc, the center of which is the pixel point to be detected, the intensity-strongly-variable points can be selected as the crossing points between arc curve and corner point edge. The connection line's orientation between above-mentoined crossing points and the center point are taken as the candidate edge orientation of the corner points. To restrain the disturbance created by the little variety in the image's partial detail, we define a partial average function,

$$g(i,j)=\frac{1}{N_{mO(i,j,M_g)}}\sum f(i,j) \tag{1}$$

where $O(i, j, M_g)$ represents a circular area, the center of which is point (i, j) and the radius of which M_g; $f(i, j)$ is the intensity value of point (i, j), and N_m the number of the pixel points in the circular area. See also Fig. 1 (a).



(a) (b)

Fig. 1 corner point (a) and discrete arc (b)

Fig. 1 (b) is a kind of approximate arc curve in the discrete field, defined in the algorithm. Supposing a radius R_n, selecting a start point and coding them in turn clockwise, a series of point set can be obtained, $C_{Rn}=((i_1, j_1), (i_2, j_2),\ldots, (i_l, j_l))$. Then calculate the following sequence values in turn:

$f(i_1, j_1)$-$g(i_1, j_1)$, $f(i_2, j_2)$-$g(i_2, j_2)$,…, $f(i_l, j_l)$-$g(i_l, j_l)$.

Count the sign change times of the above-mentioned difference value N_c and note the location of the change points when the sign is changed. When N_c=2, total two crossing points between arc curve and corner point edge were obtained and marked as (i_a, j_a), (i_b, j_b). Then estimate the inclusive angle of these three points (i_a, j_a), (i,j), (i_b, j_b) (≤180°) whether is in a certain range (such as 45° ~ 135°). If it is the case, the point (i, j) is selected as candidate point and this inclusive angle value is noted, otherwise the point (i, j) is got rid as a disturbance point. The orientation of start edge of the inclusive angle is clockwise selected as the edge orientation of the candidate corner point, such as the OB edge in Fig. 1 (a). When $N\neq2$, the points are not considered.

The points selected through above steps are not all the corner points, so a corner point response function is introduced,

$$CR(i,j)=\min_{C_i}((f(i_k,j_k)-f(i,j))^2+(f(i_{k'},j_{k'})-f(i,j))^2) \tag{2}$$

where C_i is the scanning arc curve, k and k' the symmetrical point pair about the center point (i, j) of C_i, shown as Fig. 1 (b). Selecting the threshold properly, the erroneous corner points can be got rid by corner point response function.

2.2 The selection of corner points

After finding the corner points in each image, the corresponding relation of the corner points to each other between the images should be confirmed. Here, an approach of utilizing the edge orientation of the corner points to search the corresponding relation between the different corner points is proposed. Suppose a series of detected corner points in the two images f_1 and f_2 are respectively: $P=\{p_i, i=1,2,\ldots,N_1\}$, $Q=\{q_j, j=1, 2,\ldots,N_2\}$, then the difference of edge orientation between point p_i and q_i and the difference of inclusive angle are α and β respectively. A mutual correlative function (including the angle) is defined as,

$$C_f(p,q,\alpha)=\frac{1}{\delta_1\delta_2(2M+1)^2}*\sum_{\hat{x},\hat{y}=-M}^{M}[f_1(p_x+x,p_y+y)-\mu_1]*[f_2(q_x+\hat{x},q_y+\hat{y})-\mu_2] \tag{3}$$

where μ_i and δ_i are the partial averaging value and square error of $f_i(x, y)$ respectively, $(2M+1)^2$ the area of matching window,

$$x=\hat{x}\cos\alpha-\hat{y}\sin\alpha,\ \ y=\hat{x}\sin\alpha+\hat{y}\cos\alpha.$$

Only when β is larger than a threshold (here is 15^0), can C_f be judged whether it is larger than a threshold (here is 0.80). Finally, the point pair (p_i,q_i) satisfied concurrently is considered as the candidate matching point pair, thus getting the candidate matching point pair set : $CP=\{(p_i, q_j), i=1,2,\ldots,N_{cp}\}$. Then, the two point pairs (p_i, q_i) and (p_j, q_j) are selected arbitrarily in CP and the distances dp and dq between p_i and p_j and between q_i and q_j respectively are calculated. If the difference between dp and dq is smaller than a threshold given previously, then the two points are considered to be consistent with each other. At last, each point pair is judged to obtain the last correct matching point pair set: $MP=\{(u_i, v_i), i=1,2,\ldots,N_{MP}\}$, where u_i and v_i are the points in f_1 and f_2 respectively. By selecting the control points in matching set MP, the transformation parameters $\vec{a}^{\,o}$ can be get by using affine transformation model.

3. A HIERARCHICAL IMAGE REGISTRATION ALGORITHM

In the next subsections, we describe, in detail, each of the steps of the hierarchical algorithm.

3.1 Image velocity and optical flow

In a spatial coordinate system, we can describe a 3D point by its position vector. The instantaneous 3D velocity of the point is the derivative of the position vector with respect to time *t*, shown as,

$$\vec{V} = \frac{d\vec{X}}{dt} = (\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt})^T \quad (4)$$

As the 3D point moves over time, its corresponding 2D image point $(x, y)^T$ traces out a 2D path, the derivative of which with respect to time t is the image velocity,

$$\vec{\theta} = (\frac{dx}{dt}, \frac{dy}{dt})^T \quad (5)$$

The vector field of 2D velocity is commonly called the optical flow field.

3.2 Gradient-based optical flow measurement

Concretely, optical flow can be considered as instantaneous velocity field created by the pixel point with intensity which moves on the plane. Suppose $L(x, y, t)$ is the intensity of image point (x, y) at time t, $\theta_1(x, y)$ and $\theta_2(x, y)$ the horizontal velocity and vertical velocity respectively, the image point (x, y) moves to point $(x+dx, y+dy)$ at time $t+dt$, where

$$dx = \theta_1 dt, \quad dy = \theta_2 dt$$

We can assume that,

$$L(x, y, t) = L(x + dx, y + dy, t + dt) \quad (6)$$

If considering the image intensity to be a continuous variable function relative to its location and time, the right side of above formula can be approximately replaced by a first-order Taylor series expansion,

$$\vec{L}_s^T \vec{\theta}(x, y) + L_t = 0 \quad (7)$$

where

$$L_s = (L_x, L_y)^T, \quad \vec{\theta}(x, y) = (\theta_1, \theta_2)$$

L_x, L_y and L_t are respectively the spatial and temporal partial derivatives of image intensity. This is the gradient constraint.

3.3 Registration based on affine model

Some researchers indicated that when the distance between the scene and the camera is larger, it is usually possible to approximate the motion of the scene using a single affine transformation (Heeger, 1992). Here, an affine flow field can be expressed, shown as formula (8),

$$\vec{\theta}(x, y) = M(x, y)\vec{a} \quad (8)$$

where

$$\vec{a} = (a_1, a_2, a_3, a_4, a_5, a_6)^T,$$

$$M(x, y) = \begin{pmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{pmatrix}$$

Combining (7) and (8) we have:

$$\vec{L}_s^T M\vec{a} + L_t = 0 \quad (9)$$

We want to estimate out a single set of affine transformation parameters $\vec{a}$ which will be approximately valid for the whole image. Towards this goal, consider the square error in (9) summed over the image for a given $\vec{a}$ which is,

$$E(\vec{a}) = \sum_{x,y} (\vec{L}_s^T M\vec{a} + L_t)^2 \quad (10)$$

To minimize $E(\vec{a})$ we take derivative with respect to $E(\vec{a})$, which gives the solution,

$$\vec{a} = R^{-1}\vec{S} \quad (11)$$

where

$$R = \sum_{x,y} M^T \vec{L}_s \vec{L}_s^T M, \quad \vec{S} = -\sum_{x,y} M^T \vec{L}_s L_t$$

A general 2D affine transformation, which maps a point (x_1, y_1) of the first image to a point (x_2, y_2) of the second image, can be represented by,

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{13} & p_{14} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \end{pmatrix} \quad (12)$$

Considering the optical flow model in (5) and (8), we have (dt=1),

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 + a_2 & a_3 \\ a_5 & 1 + a_6 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_4 \end{pmatrix} \quad (13)$$

Comparing (12) with (13), we can see that the parameters of a general 2D affine transformation can be completely determined by an affine flow vector $\vec{a}$ obtained from the previous procedures.

3.4 Iterative and multi-scale estimation

Here, we use a simple and efficient coarse–to–fine hierarchical approach because of its good effect and large robusticity. First, we estimate the optical flow using a coarse scale and use estimated motion to register roughly the images. Following this, the estimation of residual motion is undertaken at a finer level. In fact, the process of registration is the one of estimating the affine transformation parameters. For a given pair of images, we select one image as the reference one and other as the input one, warping the input image to register with the reference image. Giving an initial estimation,

$$\vec{a}^0 = (a_1, a_2, a_3, a_4, a_5, a_6)^T$$

and an improved parameter $\vec{a}$ for the affine flow parameters, we obtain a warped image $h(x,y,t)$,

$$h(x, y, t) = W[L(x, y, t), \vec{\theta}^0] = L(x + \theta_1^0 dt, y + \theta_2^0 dt, dt) \quad (14)$$

Applying the gradient constraint on the warped image, we have ,

$$\vec{h}_s^T \Delta\vec{\theta} + h_t = 0 \quad (15)$$

where $\vec{h}_s$ and h_t are calculated using data obtained from the last step in the iterative process similar to (11), we have,

$$\triangle \vec{a} = \vec{a}^1 - \vec{a}^0 = R_1^{-1} \vec{S}_1 \quad (16)$$

where

$$R_1 = \sum_{x,y} M^T \vec{h}_s \vec{h}_s^T M, \quad \vec{S}_1 = -\sum_{x,y} M^T \vec{h}_s h_t$$

Thus, we can refine the parameters of the affine model by iterative method using formula (17),

$$\vec{a}^1 = \vec{a}^0 + \triangle \vec{a} = \vec{a}^0 + R_1^{-1} \vec{S}_1 \quad (17)$$

4. EXPERIMENTAL RESULTS

Conducting simulation in Matlab 6.0, the above algorithm is applied to the two image data shot by the same digital camera in different times. The size of the two source images are 512×512. The images are decomposed using Harr wavelet (the decomposition levels are 3. The accurate registration parameters are achieved by using the method specified above. In our experiment, we also compared our approach with that we believed to be a typical example of a feature-based approach, the one introduced in (Li, 1995). To compare the registration accuracy between the two approaches, we use the root mean square error (*RMSE*) and the correlation *R* between the overlapping areas of the registered image pairs, which are defined as

$$RMSE = \sqrt{\sum_{x,y\in\Re}(I_r(x,y)-I_w(x,y))^2 / N)} \quad (18)$$

and

$$R = \frac{2\sum_{x,y\in R1} I_r(x,y)I_w(x,y)}{\sum_{x,y\in R1}(I_r^2(x,y)+I_w^2(x,y))} \quad (19)$$

where the I_r is the original reference image, I_w the warped image, and $R1$ the overlapping area between the reference image and the warped target image. In (18) N is the total number of pixels in this overlapping area.

In Fig. 2 (a) and (b) are the source images, (c) and (d) the results of contour-based method and our method respectively.

Table 1 shows the accuracy of the registration in terms of the *RMSE* and *R*. The results show that in the examples we tested, our image registration approach can provide a more accurate estimation than the one introduced in (Li, 1995). Our image registration approach illustrates that combining of intensity-based and feature-based methods can enhance performance.

5. CONCLUSIONS

In this paper, we propose a hybrid registration algorithm that combines the intensity-based method with the feature-based method. This hybrid registration algorithm can solve the aliasing problems created by using intensity-based method alone, and register the images where exists larger offset (proportion, rotation, motion). Using the result of the feature-based method as the starting point for the intensity-based multi-scale recursive approach is a good way to employ the advantages of the both methods. Because the basic image features are used and the last estimation process is a recursive optimized process, the hybrid method is more accurate than the method based on feature alone and its robusticity is better. The experiment shows this algorithm is suitable for registering DCIF images and some multi-spectral images. The more the number of decomposition level is, the less the information included in the top level is, so it can make the registration inaccurate because the little information for registration. The optimal number of decomposition level is 3 or 4, the registrable accuracy is higher at this time. There is a limitation of our approach. The overlapping area of the two source images should not be too small, otherwise, false matching may occur due to the lack of common features. We think that combining several registration methods to avoid the limitation of the individual methods is a good way to obtain better registration performance.

REFERENCES

W.K. Pratt. Correlation techniques of image registration. *IEEE Trans. Aerospace Electronic Systems* AES-10 (3) (1974) 353-358.

A. Goshtasby, G.C. Stockman, C.V. Page. A region-based approach to digital image registration with subpixel accuracy. *IEEE Trans. Geosci, Remote*

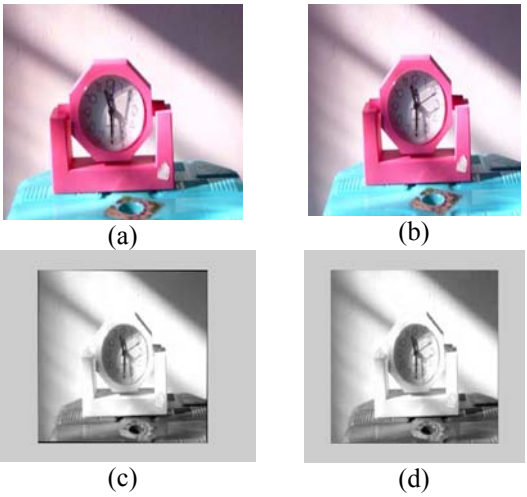
Peter Rogelj. Point similarity measures for non-rigid registration of multi-modal data. *Computer Vision and Image Uderstanding* 92 (2003) 112-140.

D.J. Heeger, A.D. Jepson. Subspace methods for recovering rigid motion: algorithm and implementation. *Int. J. Computer Vision* 7 (1992) 95-117.

H. Li. B.S. Manjunath. S.K. Mitra. A controur-based approach to multisensor image rergistration. *IEEE Trans. Image Process.* 4 (3) (1995) 320 334.

Table 1 The results of the two methods

	Contour	Our
RMSE	25.43	24.69
R	0.9512	0.9627
T_x, T_y	74.2, 4.3	75.0, 5.1
Rotation change	3.5^0	6.5^0



(a) (b) (c) (d)

Fig. 2. Images used for experiment and the results of registration figure