Image Registration Based on Both Feature and Intensity Matching

Yao Jianchao

Signal Processing Laboratory, DSO National Laboratories,
20 Science Park Drive, Singapore 118230
yjiancha@dso.org.sg

Abstract

Image registration is one of the most important tasks in image processing. The algorithms of image registration are classified into two categories: the feature-based matching and intensity-based matching. Each of them has its strength and weakness. In this paper, by combining these two techniques together, we developed a new algorithm for image registration. The algorithm utilises a parametric projective model accounting for geometrical variation and a polynomial model with a small number of polynomial coefficients explicating the smooth spatially varying illumination variation. The initial projective model parameters are first estimated by using feature-based approach. Subsequently, the coefficients of the illumination model are determined simultaneously with the projective transformation parameters through the process of intensity matching. The experimental results demonstrated the algorithm is of robustness, efficiency and accuracy.

1 Introduction

Image registration (or image matching) finds a variety of application in image processing, such as target tracking, motion analysis, objective recognition, automated visual inspection and video mosaic. A wide repertoire of image registration algorithms [1] have been reported in literature to improve the accuracy, generality, robustness and speed of image matching. They can be classified into two categories: the feature-based matching approach and intensity-based matching approach. The feature-based matching approach requires reliable feature extraction as well as robust feature correspondence to overcome missing feature and outlier problem due to partial occlusion and multiple motion. Its main advantage is the relative robustness against illumination change. It is also easy to adapt to handle non-rigid scene. However, its accuracy is up to certain amount of precision, and more importantly, the algorithm is more sensitive to the error of feature extraction and matching, for only a small portion of available image intensity information is used. On the other hand, the intensity based matching approach is mainly based on the SSD (Sum of Squared Difference) formulation [2]. Instead of requiring feature extraction or direct correspondence between two sets of features, it makes direct and complete use of all available image intensity information, thereby increasing accuracy and robustness of estimation. However, this approach suffers from a number of limitations. First of all, it is more sensitive to illumination changes than the feature-based approach. Secondly, it tends to converge to the local minima, in particular when the initial values of model are inaccurate. Thirdly, the conventional SSD-based formulation is not robust against occlusion or multiple motion.

In order to overcome aforementioned problems associated with both feature-based and intensity-based registration algorithms, we proposed a new framework for accurate image registration. The characteristic of the framework is to combine both intensity and feature matching technique together. In this framework, we utilised a rich model representing variation in both illumination and geometry. The model consists of two parts: one is a parametric projective model accounting for geometrical variation and the other a polynomial model with a small number of polynomial coefficients explicating the smooth spatially varying illumination variation. In terms of modelling image variability caused by motion, deformation and illumination, this model is comprehensive and generic. It is however quite difficult to estimate the model parameters pertaining to both projective transformation and illumination change simultaneously, due to its non-linear optimisation problem in nature. Although existing approach [3][4] that utilised the multiple-model and multi-resolution worked well for quite large class of image pair, they sometimes failed, in particular when geometrical variation between two frames is large. Instead of estimating the model parameters solely based on intensity match, in our approach, the initial value of parametric projective model are first estimated by using feature-based approach. Subsequently, the coefficients of the illumination model are determined simultaneously with the projective transformation parameters through the process of intensity matching. Because the robust estimation mechanism was used both in feature matching and intensity matching steps, the algorithm has the capability of aligning two images having partial occlusion and multiple motion. The experimental results demonstrated the algorithm is of robustness, efficiency and accuracy.

2 The formulation for image registration

In this paper, a generalised dynamic image model proposed in [8] was used. The model assumes the illumination multiplication and bias factor (α and β) to be functions of location (*x,y*). In most cases, since these two illumination factors are slowly varying functions of location, thus they can be well approximated by low-order polynomial of (*x,y*) as follows.

$$g(x^{'}, y^{'}) = \alpha(x, y) \times f(x, y) + \beta(x, y) \qquad (1)$$

Where $\alpha(x, y)$ and $\beta(x, y)$ are the low-order polynomial functions with the coefficients represented by $\vec{\alpha} = (\alpha_0, \alpha_1, \cdots \alpha_{p-1})$ and $\vec{\beta} = (\beta_0, \beta_1, \cdots \beta_{p-1})$ respectively. In our implementation, for simplicity, we employed a bilinear model for the illumination multiplication function and a constant function for the illumination bias. In this simple case, $\vec{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$ and $\vec{\beta} = (\beta_0)$. It can be readily extended to a more complicated case. The coefficients of these two polynomial functions can be determined simultaneously with the geometric transformation parameters.

The geometric transformation in the image matching for planar objects can be strictly represented by a projective transformation. For projective transformation, the position relationship between a pair of corresponding points can be written as:

$$x_i^{'} = \frac{m_o x_i + m_1 y_i + m_2}{m_6 x_i + m_7 y_i + 1}, \quad y_i^{'} = \frac{m_3 x_i + m_4 y_i + m_5}{m_6 x_i + m_7 y_i + 1}$$

(2)

Where $\vec{m} = \{m_0, m_1, \cdots m_7\}$ is the parameter vector for a projective model.

The error function between two frames after projective and illumination transformation will be represented as:

$$\sum_i | g(x_i^{'}, y_i^{'}) - \alpha(x_i, y_i) f(x_i, y_i) - \beta(x_i, y_i) |^2$$

(3)

The alignment of two images is to find such optimal parameters $\vec{m}, \vec{\alpha}, \vec{\beta}$ that minimise a weighted sum of errors, i.e.,

$$E = \sum_i w_i (g(x_i^{'}, y_i^{'}) - \alpha(x_i, y_i) f(x_i, y_i) - \beta(x_i, y_i))^2$$

(4)

In our framework of image registration, the initial estimate of projective parameters is obtained by feature-based match approach, it is then refined with illumination model parameters via Levenberg-Marquardt algorithm based on intensity-based match strategy.

3. Estimation of Initial values of Projective Parameters Based on Feature matching

3.1 Feature extraction and tracking

The feature is defined as the centre of the region having high spatial gradient and determined in terms of its "trackability"[6]. After extracting the features in the first image, their corresponding points in the second image are estimated by tracking algorithm based on translational motion model, proposed by Kanade [6]. In order to deal with an image pair with large baseline, multiresolution technique was incorporated into our tracking algorithm. The output of this step is a set of feature correspondences.

3.2 Outlier detection based on Least Median Squares

Being a good correspondence must satisfy some constraints. The epipolar geometry is the basic constraint that arises from the existence of two viewpoints. It says that a corresponding point of each one in the first retina must lies on its epipolar line. By measuring the distances of a given correspondence to their corresponding epipolar lines, it is possible for us to determine how good the correspondence is. It is however, the prerequisite of utilising above idea to distinguish between a good correspondence or a bad correspondence, termed as "outlier" on our context, is the accurate estimation of epipoler geometry from the initial correspondence set that might contains possible false matches.

Using robust regression method, called least median of squares (LMS), for outlier detection, and further for computation of fundamental matrices is now standard stuff. Interested reader can see references [7][8].

3.3 Initial estimate of projective parameters based on good correspondence set

After outlier removal, the good correspondence set is denoted by $\{(x_i, y_i)(x_i^{'}, y_i^{'})\}(1 \le i \le n)$. According to equation (4), every correspondence gives us two linear equations about unknown parameters $\vec{m} = \{m_0, m_1, \cdots m_7\}$. If the number of correspondences is more than 4, the initial projective parameter values can be determined via Moore-Penrose pseudo-inverse approach.

The accuracy of estimated projective model parameters is limited, since only a small portion of intensity information was utilised. The error $\vec{m}$ is much sensitive to that of image co-ordinates of correspondences, in particular there is large geometric variation between two frames. It needs to be further refined via intensity matching. As for the initial values of illumination model parameters $\vec{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$ and $\vec{\beta} = (\beta_0)$, it is well justified from physical appreciation that (1,0,0) and (0) can be acted as their initial value respectively.

4. Parameter Refinement Based on Intensity Matching

4.1 Robust estimation via dynamic weighting

Since images may contain partial occlusion, multiple motion and perturbation noise, robust estimation technique was used in the paper. In robust estimation, the quadratic function of residue used in least-squares estimation is replaced by a ρ-function, which assigns small weights for the constraint with larger residues. There are a few member of ρ-functions used in computer vision, we used Lorentzion function in the paper, which is given as follows:

$$\rho_{LO}(x,\sigma) = \log(1+\frac{x^2}{2\sigma^2})$$

Where x is the residue of data constraint and σ is the scale parameter. To use robust estimation in our minimisation framework, we can simple replace the quadratic function given in equation (4) by above ρ-function, This yields the following new objective function

$$E(\vec{m},\vec{\alpha},\vec{\beta}) = \sum_i w_i \rho_{LO}(r_i,\sigma) =$$

$$\sum_i w_i \rho_{LO}(g(x_i^{'},y_i^{'}) - \alpha(x_i,y_i) f(x_i,y_i) - \beta(x_i,y_i),\sigma)$$

(5)

4.2 Optimisation Algorithm

Once the initial estimate of both projective and illumination model parameter were given, we can solve the minimisation problem of (5) by using Levenberg-Marquardt algorithm. The algorithm requires computation of the partial derivatives of r_i defined in (5) with respect to projective model parameters $\{m_0, m_1, \cdots m_7\}$ and illumination parameters $\vec{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$ and $\vec{\beta} = (\beta_0)$. These are straightforward to compute. For example,

$$\frac{\partial r_i}{\partial m_0} = \frac{x_i}{D_i}\frac{\partial g}{\partial x^{'}}, \cdots\cdots, \frac{\partial r_i}{\partial m_7} = -\frac{y_i}{D_i}\left(x_i^{'}\frac{\partial g}{\partial x^{'}} + y_i^{'}\frac{\partial g}{\partial y^{'}}\right)$$

$$\frac{\partial r_i}{\partial \alpha_0} = -f(x_i,y_i), \frac{\partial r_i}{\partial \alpha_1} = -x_i f(x_i,y_i), \cdots \frac{\partial r_i}{\partial \beta_0} = -1$$

(6)

Where $D_i = m_6 x_i + m_7 y_i$ and $(\frac{\partial g}{\partial x^{'}}, \frac{\partial g}{\partial y^{'}})$ is the image intensity gradient of g at $(x_i^{'}, y_i^{'})$. From these partial derivatives, we constructed the following iterative procedure based on the Levenberg-Marquardt algorithm:

1. Initialise model parameter $\vec{c}_0 = (\vec{m}_0, \vec{\alpha}_0, \vec{\beta}_0)$ based on feature based matching. The initial σ was selected to be larger enough so that the robust estimation is very much like the least-squares estimators at beginning. Set the iteration index k=0;
2. Compute the residues r_i and the associated gradient vector $\vec{gad}_i = \frac{\partial r_i}{\partial \vec{c}}$ based on equation (6).
3. Compute the weight τ_i associated with each data constraint by $\tau_i = w_i \frac{\dot{\rho}(r_i)}{r_i}$
4. Form the weighted Hessian matrix and the weighted gradient vector $H = \sum_i \tau_i \vec{gad}\,\vec{gad}^T$, $\vec{b} = -\sum_i \tau_i r_i \vec{gad}$.
5. Update the motion parameter estimate $\vec{c}$ by an amount $\Delta\vec{c} = (H+\lambda I)^{-1}\vec{b}$, where λ is a time-varying stabilisation parameter.
6. Update the scale parameter $\sigma = \sqrt{\sum_i \tau_i r_i^2 / \sum_i \tau_i}$
7. Set $\vec{c}_{k+1} = \vec{c}_k + \Delta\vec{c}$ and k=k+1; if $\Delta\vec{c} = 0$, stop ; else go back to step 2

The advantage of using Levenberg-Marquardt over straightforward gradient descent is that it converges in less iteration.

5 Experimental Results

Figure 1 shows an example of alignment of two images taken by aerial digital camera. The images were taken under different illumination condition. In addtion, geometric variation between two images is quite large. For such kind of images, it is quite difficult to align them accurately by using conventional multiple model and multi-resolution algorithm, since initial parameter values of a model are far wary from zero. By using our algorithm, the registration result is shown in figure 1 (c). Although these two images contain large variation in geometry and illumination, the algorithm aligns them precisely.

Figure 2 show an example of alignment of two SPOT satellite images. The variability in these two images of the captured scene are extremely large, caused by the variation of incident angle of camera, variation in illumination and cloud occlusion. In addition, the quality of images, especially for first one, is very bad. Even for such an image pair, we are still able to get a satisfactory registration result by utilising our algorithm. As shown in Figure 2, there are three reference cross signs in the first image (a). The figure (b) shows accurate projective alignment under both cloud occulusion and illumination change by illustrating the estimated cross sign positions.

6 Conclusion

In the paper, a new image registration algorithm was developed, which has the following characteristics: (1). Combining the feature based matching technique for estimating the initial value of projective model makes the algorithm successful in dealing with large geometric variation. (2). The robust estimation mechanism was incorporated into the algorithm, thus allowing partial occlusion and multiple motion; (3) Illumination model based on low-order polynomial functions was used in the our approach enhances the robustness of the algorithm against illumination changes. The effectiveness of the algorithm was verified by an intensive experiment on a great number of real images.

6 References

[1]. L.G Brown, "A survey of image registration techniques", ACM Computing Surveys, Vol. 24, No. 4, pp. 325-376, 1992

[2]. P.Anandan, "A computational framework and an algorithm for the measurement of visual motion", Int. J. Computer Vision, Vol. 2, No. 3, pp. 283-310, 1898

[3]. M.J.Black and A.D. Jepson , (1998), "EigenTracking: Robust Matching and Tracking of Articulated Objects Using A View-based Representation", Int. J. Computer Vision, Vol 26, No. 1, pp. 63-84.

[4]. Gregory D. Hager, Peter N.Belhumeur (1998), “Efficient Region Tracking with Parametric Models of Geometry and Illumination”, IEEE trans. PAMI, vol. 20, No. 10, pp 1025-1039.

[5]. R. Szeliski and J. Coughlan, "Spline-based image registration", Int. J Computer Vision, Vol. 22, No.3, pp. 199-218, 1997.

[6]. Carlo Tomasi and T. Kanade, Shape and motion from Image streams: A Factorization method-2, point features in 3D motion, Technical Report CMU-CS-91-105, CMU,Pittsburgh, PA, Jan. 1991

[7]. R.Deriche, Z,Zhang, Q.T.Luong and O.Faugeras, “Robust recovery of the epipolar geometry for an uncalibrated stereo rig”, ECCV’94 567~576.

[8]. S. Negahdaripour," Rovised definition of optical flow : integration of radiometric and geometric cues for dynamic scene analysis." , IEEE Tran. Patt. Anal. Mach. Intel. Vol. 20, No. 9, pp 961~979, 1998

(a) (b) (c)

Figure 1. (a)(b) The two original images; (c) The registration result of (a) and (b)

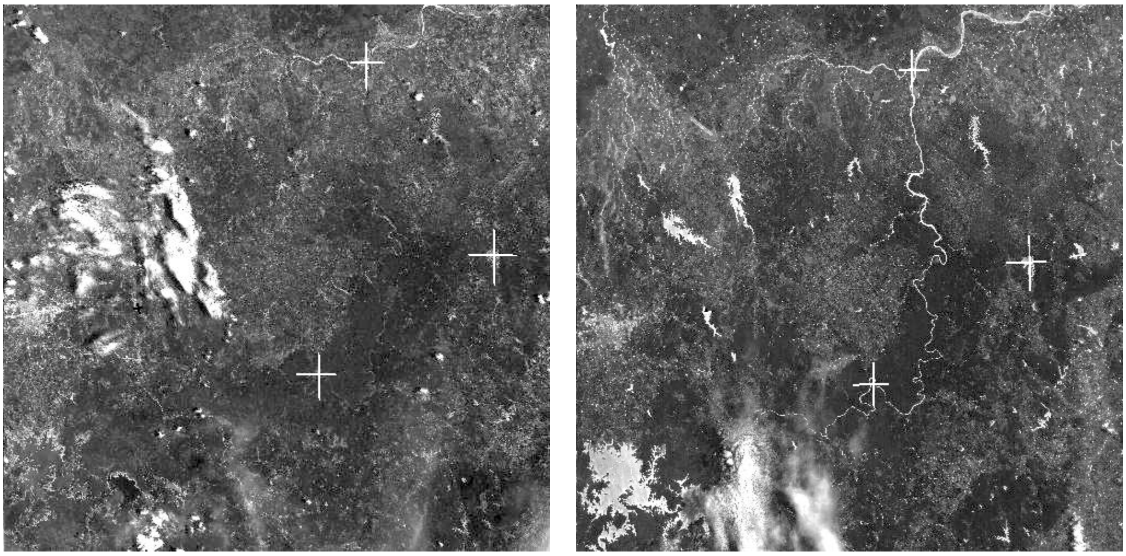

Figure 2. (a) The first image with three reference cross signs; (b) The estimated cross sign positions on second image by using estimated projective model