A novel point matching method for stereovision measurement using RANSAC affine transformation

Naiguang Lu, Peng Sun, Wenyi Deng, Lianqing Zhu, Xiaoping Lou

School of Optoelectronic Information & Telecommunication Engineering, Beijing Information Science & Technology University, Beijing 100192, P. R. China
Tel.: + 86-10-82426918 Fax: + 86-10-82426915 E-mail: nglv2002@yahoo.com.cn

Abstract

A binocular point matching method using affine transformation is presented in this paper to deal with matching ambiguities. The epipolar geometry is used to determine all the possible matching pairs to form an initial correspondence data set. Then, an affine registration model with four parameters that is invariant to scaling, rotation and translation is built using the Random Sample Consensus (RANSAC) method to describe the coordinate transformation between the two members of correspondences. Finally, correspondences are picked out using minimal nearest neighbor distances based on the geometric similarity between the right image points and the transformed left ones. The proposed method is applied to measure the profile of a 3.5m parabolic reflector of an inflatable antenna and proved to be able to handle the extra or missing point conditions coursed by occlusion and sheltering. Satisfactory results are obtained with high correct rate for most image pairs despite the significantly different viewpoints, which indicates its effectiveness and application feasibility in automatic stereovision measurement field.

Keywords: Stereo vision measurement, point matching, RANSAC, affine transformation

1. Introduction

In stereovision measurement, 3D point coordinates are computed from 2D matching points in different images. Point matching, which determines the correspondence relationship, is the key step before 3D reconstruction and plays an important part in the automation of stereovision measurement. Matching ambiguities arise in the measurement for the reasons:

1. many pictures are taken in different orientations and locations around the object in order to improve the accuracy of 3D computation. Many image points’ correspondences are invisible because of occlusion and sheltering. That is the extra/missing noise matching conditions;

2. accurate 3D coordinates as well as camera orientation and distortion parameters are obtained after point matching procedure. Consequently, accurate and reliable epipolar geometry used widely in point matching of stereovision measurement is unavailable.

Image transformation has found its wide use in point matching and object recognition. A cluster method to match point patterns using a special kind of affine transformation which is invariant to scaling, rotation and translation is proposed in [1]. A point pair with the most matching supports is picked out as a correctly matching pair and the invariant quantities computed are considered as the registration parameters between the two point patterns. Point pairs that are consistent under the registration model when vectored with the selected matching pair are collected into the correspondence set. Finally, the least-square optimal registration parameters are estimated using correspondences in the set.

The robust method in [1] gives perfect results soon when matching point patterns with scaling, rotation and translation changes only, such as the fingerprint recognition problems. The same registration model is introduced to match object points in our stereovision measurement. Because stereovision point matching is much more complicated than
fingerprint recognition cases, a totally new algorithm is proposed in this paper to improve the accuracy in global transformation parameters estimation so as to enhance the matching constraint while reducing the computation complexity.

2. Principle
2.1 Building the possible correspondence set by epipolar geometry

In epipolar geometry, a point’s correspondence lies in its epipolar line in the other image, which is the mostly widely used constraint in photogrammetry. A compact image rectification algorithm is introduced in [2] which conducts orientation transformation for each image so that conjugate epipolar lines run collinear and horizontal. As a result, the search for a point’s correspondence is restricted in one-dimension to find the image point with the same vertical coordinate.

Image rectification is carried out first to reduce the complex of epipolar geometry and increase distribution similarity, which is helpful to match points through geometrical transformation between images with more points. Generally, a point’s correspondence doesn’t lie exactly in its epipolar line for the reason of distortion and inaccurately computed fundamental matrix. Therefore, a distance bound is introduced which is called the vertical coordinate difference bound for rectified images. For each point in the left image, a right point giving a distance from the epipolar line smaller than the bound is considered as one of its possible correspondences and the point pair is collected into the possible matching point pair set.

In our measurement, a bound of 10 pixels is used to detect possible matching pairs and many points even have up to 5 possible correspondences. So, a considerable number of matching pairs in the possible set are spurious ones. More severely, even a one-to-one correspondence in the set doesn’t give a correct match for the reason of occlusion and sheltering. Matching points by transformation similarity under these two kinds of ambiguities as described in [3] are the motivations of our method.

2.2 Calculating the Affine Transformation Parameters by RANSAC

The affine transformation model mentioned above is given by:

\[
q = G(p) \Rightarrow \begin{pmatrix} x_q \\ y_q \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{pmatrix} s \cos \theta & -s \sin \theta \\ s \sin \theta & s \cos \theta \end{pmatrix} \begin{pmatrix} x_p \\ y_p \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{pmatrix} s_1 & -s_2 \\ s_2 & s_1 \end{pmatrix} \begin{pmatrix} x_p \\ y_p \end{pmatrix}
\]

where \( s \) is the scaling factor, \( \theta \) the rotation angle and \( t_x, t_y \) the translations along the horizontal and vertical directions in the image respectively; \( p(x_p, y_p) \) is a point in the left image and \( q(x_q, y_q) \) is the correspondence point in the right image. The objective of this method is to acquire a global transformation model \( G(t_x, t_y, s_1, s_2) \) under which a maximum number of correct matching pairs are consistent. A point pair \( p \leftrightarrow q \) is consistent under \( G \) if the point \( G(p) \) differs from \( q \) within a given bound \( d_e \), i.e. \( \|G(p) - q\| \leq d_e \), where \( d_e \) is the allowable distortion bound in parameter estimation [1].

Martin A. Fischler and Robert C. Bolles introduced the RANSAC model fitting method and a paradigm of its applications in [4]. The RANSAC method can detect gross error data members so as to avoid their negative affects upon the traditional model fitting techniques such as least-square method. It’s reasonable to apply this method in the parameter estimation from a data set within which lots of gross matching errors exist.

For any two possible matching pairs, a transformation \( G_i(t_x, t_y, s_1, s_2) \) is derived using equation (1). Other possible matching pairs consistent with this \( G_i(t_x, t_y, s_1, s_2) \) are selected into a compatible data subset \( S_k \). The compatible subset having the maximum point pairs is picked out to fit the global affine transformation model.

Equation (1) can be rewritten as follows for the correspondences in \( S_k \):
where $p_k \leftrightarrow q_k$ is any one of the matching pairs in the compatible subset. Then the optimal least-square transformation parameters are given by:

$$r = (C^T C)^{-1} C^T t$$  

(3)

where $C = [C_1^T, C_2^T, \ldots, C_n^T]$, $t = [t_1^T, t_2^T, \ldots, t_n^T]$ ($n$ is the length of the subset). Compared with the $G_0(t_1, t_2, s_1, s_2)$, parameters $(t_1, t_2, s_1, s_2)$ in $r$ are global and more precise.

2.3 Matching points

Points are matched based on distribution similarity between the transformed left points and right ones. Each point in the left image $p$ is transformed under the registration model $G_0(t_x, t_y, s_1, s_2)$ computed by (3) and the right point $q$ that satisfies the following condition:

$$\|q - G_r(p)\| \leq d_m$$  

(4)

and is nearest to $G_0(p)$ is considered as $p$’s correspondence, where $d_m$ is the allowable variance bound in point matching. Then, $p \leftrightarrow q$ is collected into the matching set $M$.

Two bound values $d_e$ and $d_m$ are used. $d_e$ determines the length of the compatible subset $S$ and the overall applicability of the least square transformation parameters while $d_m$ determines the length of the matching set $M$ and the correct rate of the correspondences in relatively dense areas. Generally, $S$ and $M$ are not equal in set length.

2.4 Matching more points by repeating

One registration model $G_0(t_x, t_y, s_1, s_2)$ can’t find all of the matching pairs for the complication of image transformation in stereovision measurement of curves and long baseline between some of the image pairs. It’s necessary to find other affine transformation models to match more points.

After the matching set $M$ has been established, all the possible pairs with points that have been matched are excluded from the possible correspondence set. The estimation and matching procedure in 2.2 and 2.3 are carried out again to compute new model parameters and obtain more correspondences those are consistent with them.

With the decrease of the possible set, length of the compatible subset gets small. To guarantee the globalization and avoid wrong matches, the minimum length of the compatible subset is fixed to $L$. $L$ is set to 5 in this paper. It’s easily understood that with the increase of $L$, the correct rate is improved while the total number of the matches decreases. When the length of the compatible set is smaller than $L$ in one trial of the repeat, then the estimation bound $d_e$ and matching bound $d_m$ are increased correspondingly to include more matches.

3. Experiment results

In our 3D measurement, retro-reflective target points for easy identification are pasted on the surfaces and pictures are captured from different orientations. The centroids of these target spots in pictures are computed as the points to be measured. 3D coordinates of these points are computed by photogrammetry from disparities between correspondences. Two kinds of experiments are carried out to test our method, plane and curved surface point matching.

3.1 Matching points on images of a plane

In this experiment, target points are pasted on the wall and some bars fixed on it. Figure 1 (a) shows one of the image pairs. Points are not rectified here.

Figure 1 (b) shows the matching results. There are 63 points in the left image and 60 points in the right image with 59 correspondences. 55 correspondences are obtained by the algorithm within 5 seconds and all of them are correct.

2-175
3.2 Matching points on images of a curved surface
In this experiment, images for the profile measurement of a 3.5m inflatable parabolic reflector of a satellite antenna are used. 3D coordinates of the target points are used to fit a mathematical description of the surface shape [5]. Points are rectified first.

Figure 2 shows the matching results. There are 93 points in the left image and 96 points in the right image with 90 correspondences. 75 correspondences are obtained by the algorithm within 1 minute and all of them are correct.

Results of the two experiments prove the effectivity and low time consumption of our method. Registration parameters are global and computed accurately so that points are mapped onto their correspondences exactly as shown in the figures by ‘╳’. What’s more, the method is robust for occlusion and sheltering by other objects in front of the measured profile, so it has little requirement of the scene environment.

4. Conclusion

An automatic binocular method of object point matching for 3D profile measurement is proposed in this paper. Affine transformation models that are invariant to scaling, rotation and translation are computed by RANSAC from all the possible matching pairs detected by epipolar geometry. Point pairs that are consistent with these registrations are considered as matching pairs. This method is insensitive to camera distortion, orientations and the precision of fundamental matrix. What’s more, it keeps robust under extra/missing conditions and achieves low time consumption. Satisfactory results of many experiments indicate its effectivity and application feasibility in practice.

5. Acknowledgements

This project is supported by National Natural Science Foundation of China (50475176, 50675015), Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (PXM2007_014224_044674).

References