Journal of Southeast University (English Edition) Vol. 23, No. 2, pp. 190 – 195 June 2007 ISSN 1003—7985

Non-iterative image feature matching algorithm based on reference point correspondences

Zhang Weizhong[1,2] Zhang Liyan[2] Wang Xiaoping[2] Ding Zhian[2] Zhou Ling[2]

([1]College of Information Engineering, Qingdao University, Qingdao 266071, China)

([2]Research Center of CAD/CAM Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Based on the coded and non-coded targets, the targets are extracted from the images according to their size, shape and intensity etc., and thus an improved method to identify the unique identity(ID) of every coded target is put forward and the non-coded and coded targets are classified. Moreover, the gray scale centroid algorithm is applied to obtain the subpixel location of both uncoded and coded targets. The initial matching of the uncoded target correspondences between an image pair is established according to similarity and compatibility, which are based on the ID correspondences of the coded targets. The outliers in the initial matching of the uncoded target are eliminated according to three rules to finally obtain the uncoded target correspondences. Practical examples show that the algorithm is rapid, robust and is of high precision and matching ratio.

Key words: reference points detection; coded and non-coded target; subpixel; gray scale centroid; point correspondence

The search for a robust and automatic solution of the image correspondence problem, or image matching, is one of the most challenging problems in computer vision and digital photogrammetry. It has important applications value in many fields, such as movement estimation, image synthesis, object recognition and tracking, etc. In computer vision and close-range photogrammetry, especially, the results of 3-D reconstruction depend on the quality of image matching.

According to the cues used for matching, existing algorithms can be divided into two categories: area-based matching[1-2] and feature-based matching[3-4]. Compared to area-based matching, feature-based matching is not subject to infection by factors of illumination and projective changes of shape so that matching results are comparatively reliable. Feature point matching algorithms can further be classified into global methods, such as dynamic programming, relaxation, exhausted search, and local algorithms, such as greedy search, simulated annealing and randomized search.

The integration of the estimation of relative orientation between the camera and the point correspondence solution between two images was discussed in Ref. [5]. Although the epipolar geometry constraints were considered in the relative orientation estimation and correspondence solution, Ref. [5] concluded that those restrictions were not enough to prevent or to diminish the occurrences of false correspondence.

The feature-based matching algorithm in Ref. [6] is established on relaxation labeling of global optimization. The involved metric of the distance relation between primitives and the volume of matching parallelepiped (MP) based on epipolar constraints, together with the metrics such as cross-correlation coefficients, the differences of gradient and average intensity, make the correct matching ratio higher and the matching process automated. The 3-D reconstruction precision is influenced since the extraction of the feature points was made manually just with pixel accuracy in Ref. [6].

Our method uses coded and non-coded targets. Both the coded and the non-coded targets are called reference points. The coded targets are utilized for both determination of distance relations between the primitives and camera self-calibration which is not discussed in this paper. The non-coded targets are utilized for 3-D point reconstruction. The reference points are extracted automatically with subpixel accuracy. Therefore, the proposed algorithm improves the automation level of the above-mentioned algorithm by automatic detection with the reference points and 3-D point reconstruction precision.

Received 2006-10-08.

Foundation item: The National Natural Science Foundation of China (No. 50475041).

Biography: Zhang Weizhong (1963—), male, professor, zhangwz_01@163.com.

1 Automatic Coded and Non-Coded Targets Detection

Our automatic reference point detection algorithm includes mainly three steps, namely: reference point ellipse contours extracting, reference point identifying and locating, as shown in Fig. 1(d).

1.1 Extracting reference point ellipse contours

The coded and non-coded targets that we used for range photogrammetry are demonstrated in Figs. 1(b) and (c). The coded target center is a circular target surrounded by a unique code band pattern that is used to identify the target. Fig. 1 (b) shows the structure of a coded target. The "code band" is composed of bit positions at equally spaced 15 angular intervals, each of the 15 portions is 24°. Each position can be either foreground color or background color, corresponding to a binary code "1" or "0".

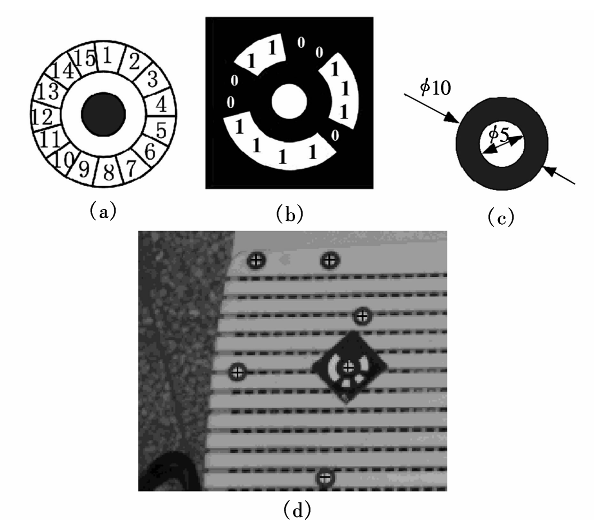


Fig. 1 Reference point. (a) Structure of coded target; (b) Illustration of coded point; (c) Non-coded point; (d) Local extracting result of the scene

The inner circle of the reference point is the target to be detected. The target of the reference point becomes an ellipse by camera imaging. The image coordinates of the ellipse center are picked up by the extracting algorithm of the image feature.

First, the image is segmented using a Canny operator. Secondly, the contour information denoting different areas is extracted. Thirdly, combining with the features of the size, shape, intensity change and position distribution of the reference points, ellipse contours are extracted. Fig. 2 shows the process of eliminating non-reference targets by using reference point features.

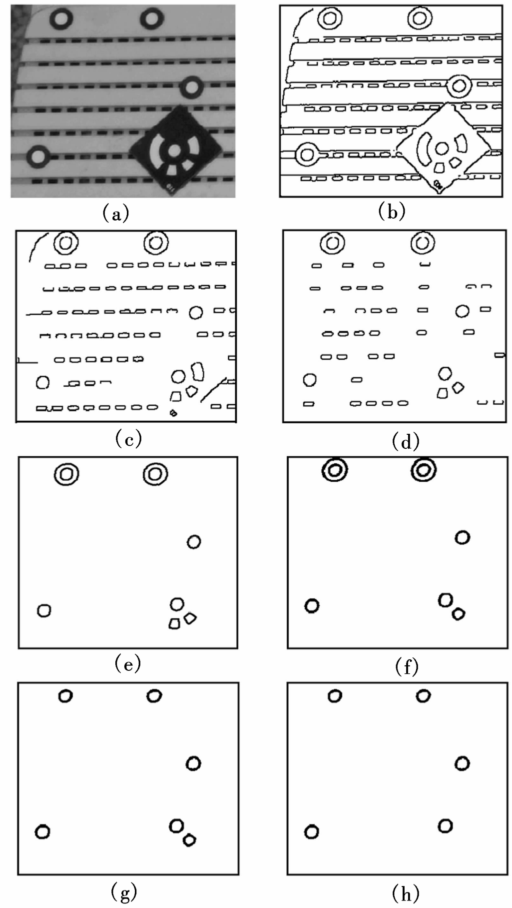


Fig. 2 The process of eliminating non-reference points by using reference point feature. (a) Local scene image; (b) Edge image; (c) The filteration result by the size rule; (d) The filteration result by the concavo-convex and close rule; (e) The filteration result by the circinal rule; (f) The fitting filteration result by least-squares; (g) The filteration result by the gray-scale rule; (h) The filteration result by the location rule

1.2 Identifying coded and non-coded targets

The identification process aims to identify connected regions of pixels that represent the boundaries of the elliptical imaged targets. This is done by identifying edge pixels within the image and then classifying connected regions of edge pixels as either coded point or non-coded point[7].

Fig. 3(a) shows an example of an ellipse (labelled D) that has been fitted to the edge pixels of the image of a circular coded target. Since the geometry of the coded target is known, ellipses that define the inner and outer boundaries of the code band (labelled C and A) and the centre of the code band (labelled B), can be computed.

The intensity value of point T_B pixel lies on the median value of window $T_C T_A$. The corresponding code for each candidate target is computed by minimizing a function of one variable. The intensity values of all

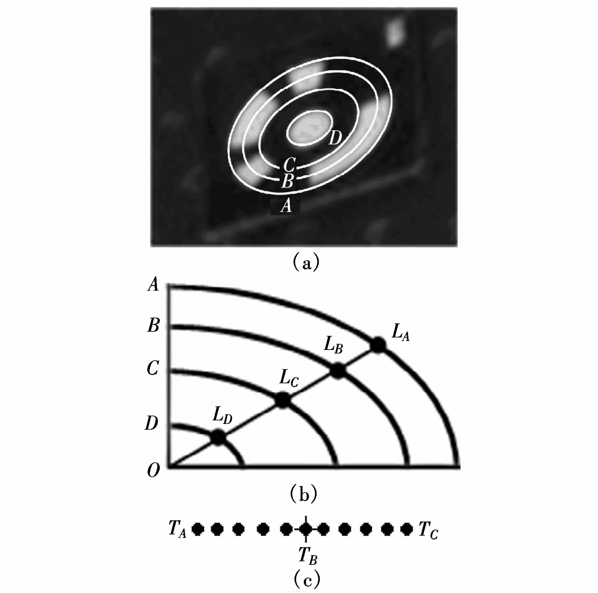


Fig. 3 Decoding coded target. (a) Coded target with fitted ellipse (labelled D), the inner and outer boundaries of the code band (labelled C and A), and the ellipse observed the code band pixel intensities on it (labelled B); (b) The relations between ellipses labelled A, B, C and D; (c) The intensity value of point T_B pixel lying on the median value of window $T_C T_A$

pixels lying on the ellipse that runs through the centre of the code band are determined. The coordinates of these pixels are normalized so that each intensity value corresponds to a position on a unit circle, rather than on an ellipse[7]. The intensity value of every pixel between the unit circle and the ellipse B is one-to-one, Starting at an angle of the unit circle is divided into 15 bit segments (for a 15-bit code), each with an angular extent of 24°. The mean intensity value corresponding to each bit segment is used to determine whether the bit segment is a "1" or a "0", consequently, the ID of the code target is obtained.

To decode the pattern, the binary code is read clockwise. Each bit is considered to be the first bit in turn. This means that there are 15 binary numbers to be considered. The smallest one among these 15 numbers is chosen to be the ID of the coded point. For instance, the 15-bit code shown in Fig. 1(b) corresponds to the binary sequences "001100111011111", "011001110111110", "110011101111100", … "110011001110111", "100110011101111" (white or empty regions represent "1"s and black or inked regions represent "0"s). Of these nine binary numbers, 001010111, has the lowest value "$001100111011111_2 = 6623$", so this code pattern is labelled as number 6623. It is also the ID of this code target.

The recognized unique identity of the coded point separates the coded targets from the non-coded ones.

The coded target number is identified by the features of the coded target. Therefore, the matching relationship between coded targets in multiple views is established easily by the number of coded targets.

1.3 Subpixel target center detection

In close-range photogrammetry, the image location of reference points are applied to camera calibration or 3-D reconstruction. Therefore, the precision of close-range photogrammetric measurements is influenced by the location algorithm of reference points. According to Ref. [8], the gray scale centroid locating method is more accurate compared to other subpixel algorithms. This method is applied to our algorithm. The target center coordinates of the reference point can be calculated by

$$\left.\begin{aligned} x_c &= \frac{\sum_j \sum_i i I_{i,j}}{\sum_j \sum_i I_{i,j}} \\ y_c &= \frac{\sum_j \sum_i j I_{i,j}}{\sum_j \sum_i I_{i,j}} \end{aligned}\right\} \tag{1}$$

where (x_c, y_c) are the target center coordinates of the reference point, and $I_{i,j}$ is the intensity value of the point (i, j) pixel.

2 Reference Point Based Matching

Now we will give a description of the algorithm of reference point based matching, which improves on the original algorithm in Ref. [6] to obtain higher efficiency and 3-D reconstruction precision. In this paper, the pair of candidates (i, j) that are the i-th and j-th feature points to be matched within two images are expressed by non-coded targets, and their neighbors (i_s, j_s) are expressed by coded targets, as shown in Fig. 4. While it is difficult to blind search within two images for corresponding points, it is easier to start from an already matched pair. To facilitate this, first, the coded targets within two images are matched by using their IDs. Secondly, the matching process of the non-coded

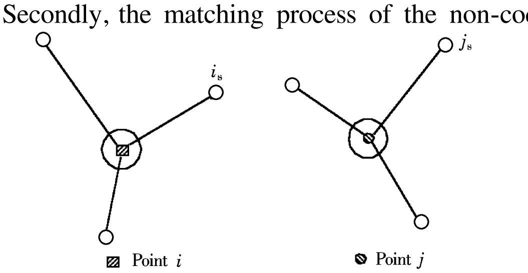


Fig. 4 The pair of candidates (i, j) and their neighbors (i_s, j_s)

targets within two images is performed based on the matching result of the coded targets.

2.1 Similarity computation

In similarity computation, there are usually various metrics for the pair of candidates (i, j), such as cross-correlation coefficient ρ_{ij}, gradient difference and average intensity difference. They are expressed as

$$S_{ij}^{cc} = \rho_{ij}$$

$$S_{ij}^{grad} = f_s(\alpha_{grad}, |g_j - g_i|)$$

$$S_{ij}^{int} = f_s(\alpha_{int}, |I_j - I_i - (\bar{I}_j - \bar{I}_i)|)$$

Then, the initial similarity for one pair (i, j) can be calculated by

$$S_{ij}^0 = S_{ij}^{cc}\, S_{ij}^{grad}\, S_{ij}^{int} \tag{2}$$

where α_x is a positive constant whose value is related to the degree of variation of S_{ij} with the distance d_{ij} between i and j, $g_{i(j)}$ is the magnitude of the gradient around $i(j)$, $I_{i(j)}$ is the average intensity around $i(j)$, and $\bar{I}_{i(j)}$ is the average intensities of the two different images.

The constraints related to epipolar geometry may be incorporated into the similarity computation, as an additional metric. The metric is calculated according to

$$S_{ij}^{epi} = f_s(\alpha_{epi}, V_{mp}(\bar{\boldsymbol{F}}, i, j)) \tag{3}$$

where V_{MP} is the volume of the matching parallelepiped for the pair (i, j), whose value is directly proportional to the distance from j to the epipolar line defined by i, and vice versa. As a result, this volume may be used as a metric associated with the attendance of epipolar geometry. For more details, see Ref. [6]. $\bar{\boldsymbol{F}}$ is the estimation of the fundamental matrix. Then the initial similarity can be rewritten as

$$S_{ij}^0 = S_{ij}^{cc}\, S_{ij}^{grad}\, S_{ij}^{int}\, S_{ij}^{epi} \tag{4}$$

2.2 Compatibility computation

In compatibility computation, there are usually metrics for the pair of candidates (i, j): distance relation between neighbors based metric C_{ij}^{dist} and volume of MP based metric C_{ij}^{epi}. They are expressed as

$$C_{ij}^{dist} = \frac{\sum_{s=1}^{N} \delta(i, j; i_s, j_s)}{1 + \mathrm{dist}(i, j; i_s, j_s)}$$

where, according to Ref. [5], $\delta(i, j; i_s, j_s) = e^{-\frac{r}{\varepsilon_r}}$, it is a Gaussian function, $r = \frac{|d(i, i_s) - d(j, j_s)|}{\mathrm{dist}(i, j; i_s, j_s)}$; $\mathrm{dist}(i, j; i_s, j_s) = \frac{d(i, i_s) + d(j, j_s)}{2}$; i_s, j_s are neighbors of the pair of candidates (i, j), and $d(\cdot)$ represents the Euclidean distance.

$$C_{ij}^{epi} = \frac{1}{1 + \alpha'_{epi} \sum_{i_s, j_s \in N} V_{MP}(\bar{\boldsymbol{F}}, i_s, j_s)}$$

where α'_{epi} is similar to α_{epi} in Eq. (3), and $V_{MP}(\bar{\boldsymbol{F}}, i_s, j_s)$ is the volume of MP of neighbors around (i, j).

Compared to the method in Ref. [6], the correspondence of the neighbors expressed by the coded targets between the two different images is achieved easily so that its corresponding volume computation of MP is easier.

The compatibility, considering distance relations and epipolar constraints, is obtained as

$$C(i, j) = C_{ij}^{dist}\, C_{ij}^{epi} \tag{5}$$

2.3 Matching solution

Compared to the method in Ref. [6], the iterative computation is not involved since the determination of relation orientation is not considered in the proposed algorithm.

First, the similarities are computed using the cross-correlation coefficients and the differences of gradient as metrics. The compatibility is computed considering the distance relations between neighbors. The initial matching of the uncoded target correspondences between an image pair is established according to similarity and compatibility, which are based on the ID correspondences of the coded targets. Secondly, the $\boldsymbol{F}$ matrix is computed using coded target correspondences matched by using the method in Ref. [9]. Then, the volume of MP is computed using the coordinates of the non-coded targets matched initially by

$$V_{MP}(\boldsymbol{F}, i, j) = |\boldsymbol{p}_l^T \boldsymbol{F} \boldsymbol{p}_r| \tag{6}$$

If

$$V_{MP}(\boldsymbol{F}, i, j) = |\boldsymbol{p}_l^T \boldsymbol{F} \boldsymbol{p}_r| < \varepsilon_{MP} \tag{7}$$

then S_{ij}^{epi} and C_{ij}^{epi} are computed, respectively, by using the volume of the MP. Otherwise, the candidate pairs of the non-coded targets are eliminated from its initial matching solution. Where ε_{MP} is the threshold of the volume of the MP. The constraints related to epipolar geometry based on the volume of the MP may be incorporated into the similarity computation by Eq. (4), as an additional metric and the similarity is recomputed. The constraints related to epipolar geometry based on the volume of the MP may be incorporated into the compatibility computation by Eq. (5), as an additional metric and the compatibility is recomputed. So, the new initial matching solution of the uncoded target is obtained by the similarity and the compatibility. Finally, the outliers in the initial matching of the uncoded target are eliminated according to three rules to obtain the uncoded target correspondences.

2.4 Elimination of false matching

Once the initial solution by similarity and compatibility is obtained, the false matching can be eliminated by three rules as follows:

1) Similarity rule One threshold ε_s is defined, if $S_{ij} < \varepsilon_s$, the pair (i, j) will be excluded from the solution.

2) Ambiguity rule Ambiguity is another criterion that may be used to exclude some error matches. It is expressed by non-ambiguity factor[5] F_i, which is defined as $F_i = 1 - p^{(2nd)}/p^{(1st)}$, where (1st) and (2nd) refer to the biggest and the second biggest similarity for point i, respectively. One threshold ε_{NAF} is defined, if $F_i < \varepsilon_{NAF}$, the pair(i, j) will be excluded from the solution.

3) Triangulation error rule Triangulation error D_{tri} is the distance between vectors $\boldsymbol{P}_l$ and $\boldsymbol{P}_r$ being in the left and right images, respectively, as in Fig. 5. $\boldsymbol{p}_l$ and $\boldsymbol{p}_r$ are the image points of the 3-D point $\boldsymbol{P}$ in the left and right images, respectively. $\boldsymbol{P}'$ is a 3-D point which is a solution by using image points $\boldsymbol{p}_l, \boldsymbol{p}_r$. ε_T is the threshold of D_{tri}. If $D_{tri} > \varepsilon_T$, the pair (i, j) will be excluded from the solution.

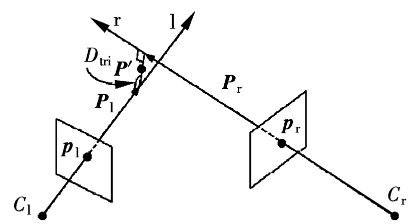


Fig. 5 Triangulation and measure error D_{tri}

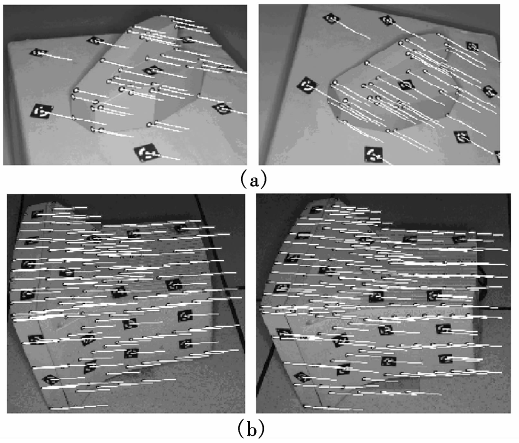
3 Experimental Results

The proposed algorithm was implemented by VC++. The real images, with 4 246 × 2 848 resolution, were obtained from a Nikon digital camera.

The multiple images are used to test our algorithm. Experiments show that the average percentage of correct matching is above 98.9%. The 3-D reconstruction precision obtained is lower than 0.4 mm/m since the subpixel center locating algorithm with a 0.02 pixel accuracy is utilized.

Fig. 6(a) is the epipolar line and reference points matching results of one of our tests. There are 8 pairs of coded targets and 46 pairs of non-coded targets. The ratio of the average correct matching is 99.2%.

Fig. 6(b) is the epipolar line and reference points matching results in another test. There are 28 pairs of coded targets and 185 pairs of non-coded targets in this experiment. The ratio of the average correct matching is 98.6%.



(a)

(b)

Fig. 6 The epipolar line and reference point matching of the scene

4 Conclusion

A non-iterative image feature matching algorithm based on reference point correspondences is proposed. The code and non-coded targets are utilized in the proposed algorithm to improve the algorithm in Ref. [6]. According to the rule of size, shape, intensity change and position distributing, the reference points in the images are detected automatically to avoid manually picking up the feature points in Ref. [6]. The matching of coded targets between the different images is reached easily by using the unique identity of coded targets. Based on the coded targets matched, the matching of non-coded targets is also easier. The locating precision of a reference point center is higher than that of the existing method in Ref. [6] by using the subpixel target center locating algorithm. The initial matching of the uncoded target correspondences between an image pair is established according to similarity and compatibility, which are based on the ID correspondences of the coded targets. The outliers in the initial matching of the uncoded target are eliminated according to three rules to finally obtain the uncoded target correspondences. Practical examples show that the algorithm is rapid, robust and is of a high precision and matching ratio.

References

[1] Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: theory and experiment [J]. *IEEE Trans on Pattern and Machine Intelligence*, 1994, **16**(9): 920 - 932.

[2] Stefano L D, Marchionni M, Mattoccia S. A fast area-based stereo matching algorithm[J]. *Image and Vision Computing*, 2004, **22**(12): 983 - 1005.
[3] Mount D M, Netanyahu N S, Moigne J L. Efficient algorithms for robust feature matching[J]. *Pattern Recognition*, 1999, **8**(6): 17 - 38.
[4] Carcassoni M, Hancock E R. Correspondence matching with modal clusters[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2003, **25**(12): 1609 - 1615.
[5] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Rapport de Recherche n. 2273 [R]. Shophia-Antipoles: Unité de Recherce INRIA, 1994.
[6] Galo M, Tozzi C L. Feature-point based matching: a sequential approach based on relaxation labeling and relative orientation[J]. *Journal of WSCG*, 2004, **12**(1): 1 - 8.
[7] Keith F, Anthon V, Ndimi B. An Inexpensive, automatic and accurate camera calibration method[C]//*Proceedings of the Thirteenth Annual Symposium of the Pattern Recognition Association of South Africa*. Cape Town, 2002: 1205 - 1211.
[8] Xu J, Fang Z P, Malcolm A A, et al. A robust close-range photogrammetric system for industrial metrology[C]//*Proc of the 7th International Conference on Control, Automation, Robotics and Vision*. Singapore, 2002: 114 - 119.
[9] Zhou Ling. Research on key technology of 3-D reconstruction based on multiple images by using one hand-held digital camera[D]. Nanjing: Research Center of CAD/CAM Engineering of Nanjing University of Aeronautics and Astronautics, 2006. (in Chinese)

一种非迭代的同名标记点图像特征匹配算法

张维中[1,2] 张丽艳[2] 王小平[2] 丁志安[2] 周 玲[2]

([1] 青岛大学信息工程学院,青岛 266071)
([2] 南京航空航天大学 CAD/CAM 工程研究中心,南京 210016)

摘要:利用编码元和非编码元,根据标记点的尺寸、形状及灰度变化等特征提取目标,然后利用非编码元与编码元的不同形状与灰度特征,提出一种改进的编码元自动身份识别方法,同时实现非编码元与编码元的分类;并利用质心定位方法抽取标记点中心位置,抽取的中心具有亚像素级.在利用编码元的身份信息实现同名编码元匹配的基础上,由相似性和相容性确定非编码元的初始匹配,通过3个准则从非编码元的初始匹配中剔除误匹配,最终得到同名非编码元的匹配结果.经实验验证,该算法速度快、匹配率高、鲁棒性好.

关键词:标记点检测;编码和非编码元;亚像素;质心定位方法;同名点匹配
中图分类号:TP391