Journal of Southeast University (English Edition) Vol. 24, No. 4, pp. 473 −477 Dec. 2008 ISSN 1003—7985

Fast estimation of fundamental matrix based on stripe constraints

Yang Ankang[1] Liu Jingnan[1] Yoshihisa Uchida[2] Yoshiyuki Uchida[2]

([1] School of Automation, Southeast University, Nanjing 210096, China)

([2] Department of Information Network Engineering, Aichi Institute of Technology, Toyoda 470-0392, Japan)

Abstract: In order to improve the performance of estimating the fundamental matrix, a key problem arising in stereo vision, a novel method based on stripe constraints is presented. In contrast to traditional methods based on algebraic least-square algorithms, the proposed approach aims to minimize a cost function that is derived from the minimum radius of the Hough transform. In a structured-light system with a particular stripe code pattern, there are linear constraints that the points with the same code are on the same surface. Using the Hough transform, the pixels with the same code map to the Hough space, and the radius of the intersections can be defined as the evaluation function in the optimization progress. The global optimum solution of the fundamental matrix can be estimated using a Levenberg-Marquardt optimization iterative process based on the Hough transform radius. Results illustrate the validity of this algorithm, and prove that this method can obtain good performance with high efficiency.

Key words: fundamental matrix; structured-light; stripe code pattern; stereo vision

The computation of the fundamental matrix existing between two views of the same scene is a common task in several applications in computer vision, including calibration and reconstruction, visual navigation, and visual servoing[1]. The fundamental matrix is important because it represents succinctly the epipolar geometry of stereo vision. Indeed, the matrix provides relationships between corresponding points in two views. Moreover, for known intrinsic camera parameters, it is possible to recover the essential matrix from the fundamental matrix and, hence, the camera motion between the views.

Several techniques have been developed for the estimation of the fundamental matrix from point correspondences, such as the linear criterion, the distance to epipolar lines criterion, and the gradient criterion[2-4]. The first one is a least-squares technique minimizing the algebraic error. This approach has been proven to be sensitive to image noise, and it does not consider the fact that the rank of the fundamental matrix must be equal to 2. The other two techniques take into account rank constraints and minimize a more indicative distance, the geometric error, in the seven degrees of freedom of the fundamental matrix. These result in nonconvex optimization problems that present local solutions in addition to global ones. Hence, the solution found via numerical procedures is affected by the choice of the starting point of the minimization algorithm[5-6]. Generally, this point is chosen as the estimate provided by the linear criterion and forced to be singular by setting the smallest singular value to zero, but this choice does not guarantee to find the global minimum.

In this paper, we present a new method for the estimation of the fundamental matrix. It consists of a constrained least-squares technique in which the rank constraints on the matrix are ensured by the constraints of the scene. In this way, we impose the singularity of the matrix *a priori* instead of forcing it after the minimization procedure as in the case in the linear criterion. Our aim is to define this linear constraint as the cost function in the structured-light system for improving the performance of calculation. In order to find the linear constraints, we start by showing the structured-light system using stripe code patterns. Then, we reformulate the estimation problem so that it can be developed in the Levenberg-Marquardt optimization algorithm. In the end, we provide experimental results showing that our approach leads to a more accurate estimate of the fundamental matrix.

1 Stripe Code Structured-Light System

In a structured-light reconstruction system, as shown in Fig. 1, the camera A takes images of an object illuminated by a set of light patterns from the projector B. The shape of the object is computed from the deformation of the projected grid. In our system, the camera and the projector are both uncalibrated, and the inner parameters of the projector are constant during the process of reconstruction while the inner parameters of the camera are volatile. Under the assumption of the linear and center projection, the projector can be modeled as a virtual pinhole camera with fixed inner parameters, and the single view system can be formulated using epipolar geometry. Consequently, the fundamental matrix is the mapping relation from the projector to the camera.

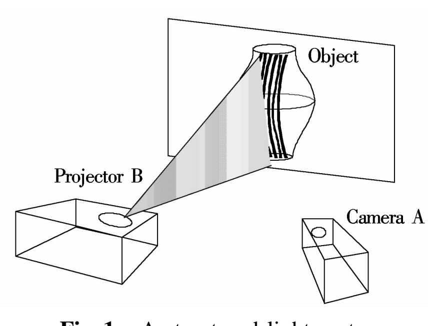

Fig. 1 A structured-light system

In order to reduce the noise of shadow in the scene, as well as extend the robustness of the image matching algorithm, we consider the appearance of a scene illuminated by a sequence of patterns using the similarity Manchester code. The system is based on the assumption that the environment background light is fixed during the project processing of the pattern images, which limits the scanning time. The assump-

Received 2008-05-28.

Biographies: Yang Ankang (1977—), male, graduate; Liu Jingnan (corresponding author), male, doctor, professor, liujn@seu.edu.cn.

Citation: Yang Ankang, Liu Jingnan, Yoshihisa Uchida, et al. Fast estimation of fundamental matrix based on stripe constraints[J]. Journal of Southeast University (English Edition), 2008, 24(4): 473 −477.

tion also provides a set of rules for designing an illumination pattern as follows: 1) The width of pattern lines is constant; therefore, the histogram of the adjacent images is similar, and the same algorithm of the image process can be valid; 2) The stripe codes are simple enough for analysis; 3) The stripe codes of adjacent pattern lines are different enough from the normal scene shadow for pattern identification; 4) The stripe boundary codes instead of the stripe codes to convey information between projector and camera are used, since at most two pixels of a given frame can be used together to infer correspondences, and the only projected feature that we can reliably identify is the boundary between two stripes. Focusing on stripe boundaries has several advantages. For example, if the stripes on either side of a stripe boundary can each be assigned n different codes over time, then nearly n^2 distinct stripe boundaries can be identified in a camera image.

Each pattern is composed of four frame stripe code images. A stripe boundary code involves assigning a binary (black or white) to each pixel in time, such that each stripe boundary has a unique code (consisting of the black/white illumination history on both sides of the boundary) over the sequence of four frames. Projecting the sequence of pattern images on the object in the scene, we can capture the reflection pattern of the object from another view. Furthermore, we can obtain the sequence of structured-light images with stripe code information. Through scanning the images, we can filter out the stripe boundary code points set and every space point on the project pattern stripe boundaries can be uniquely labeled. The stripe codes can also introduce linear constraints in the current scene; the constraints are that points with the same stripe code on the projective plane are collinear. From the above linear constraints, a new norm of optimization of the fundamental matrix estimation is deduced, which is discussed in section 3.

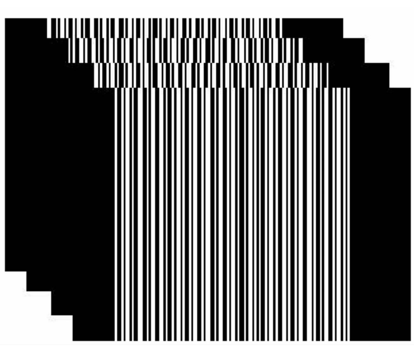

Fig. 2 Stripe code pattern images

In order to generate such patterns, we form an analogy between stripe boundary codes and build a jump list in an n-dimensional state space, for searching a maximal-length list that satisfies several conditions. In an n-dimensional state space, there are n^2 directed jumps. Using commutative algebra, we design the algorithm for the generation of the code as follows:

1) All adjacent state bits are inversed, and their list is $P_1 = \{i, \varepsilon_n\varepsilon_{n-1}\cdots\varepsilon_2\varepsilon_1\}$. The number of edges with no direction is 2^{n-1}. Put the entire list into an available set.

2) Only one bit of state is equal, and the list is $P_2 = \{i, \varepsilon_n\ldots\varepsilon_{k+1}\varepsilon_{k-1}\ldots\varepsilon_n\}(n \geqslant k \geqslant 1)$. This is a cube with $n \times 2^{n-1}$ edges. Since there is one equal bit in the code, the pattern has one invisible stripe boundary. If the adjacent pattern code has an invisible boundary at a same bit, the pattern cannot be used. Otherwise, we put the pattern code in the available set.

3) Increase the same bit number and repeat the above operation until the same bit number $k \geqslant \log 2n$.

4) In the available set, we filter the pattern code with the selection rule. For example, in the four state spaces we find about 55 patterns and 110 stripe codes. Consequently, in this paper we use stripe codes as shown in Fig. 2.

2 Basic Algebraic Estimation

Suppose that a plane ξ is imaged by two cameras at different angles. Let π and π' be images of the plane from the two cameras. If the cameras satisfy the pinhole model, the π-to-π' image transformation can be described by a planar homography, a non-singular 3×3 projective matrix. Given a pair of homogeneous coordinates of pixels $\boldsymbol{p} = \{x, y, 1\}^{\mathrm{T}}$ in π and $\boldsymbol{p}' = \{x', y', 1\}^{\mathrm{T}}$ in π' that correspond to the same point on plane ξ, there exists a planar homography $\boldsymbol{F}$ between the two image points.

$$\boldsymbol{pFp}' = \boldsymbol{0} \tag{1}$$

where

$$\boldsymbol{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

If we let

$$\boldsymbol{\tau} = \{f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33}\}^{\mathrm{T}} \tag{2}$$

be the vector of parameters, $\boldsymbol{x} = \{m_1, m_2, m_1', m_2'\}^{\mathrm{T}}$ be the vector of variables, and

$$\Gamma(\boldsymbol{x}) = \{m_1 m_1', m_2 m_1', m_1', m_1 m_2', m_2 m_2', m_2', m_1, m_2, 1\}^{\mathrm{T}}$$

be the vector of transformed variables, then Eq. (1) can be rewritten as

$$\boldsymbol{\tau}^{\mathrm{T}} \Gamma(\boldsymbol{x}) = 0 \tag{3}$$

which is the epipolar equation that we exploit to design a new method for estimating the fundamental matrix with a set of corresponding points.

Estimating the fundamental matrix is based on the use of cost functions measuring the extent to which the data, and candidate estimates fail to satisfy the epipolar equation (3). First, the rank-2 constraint is set aside for simplicity. Then, given a set of data $\boldsymbol{x}_i, \ldots, \boldsymbol{x}_n$ and a cost function $J = J(\boldsymbol{\tau}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, a corresponding estimate $\hat{\boldsymbol{\tau}}$ is defined as the parameter which minimizes J:

$$J(\hat{\boldsymbol{\tau}}) = \min_{\tau \neq 0} J(\boldsymbol{\tau}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \tag{4}$$

Eq. (1) does not change if $\boldsymbol{\tau}$ is multiplied by a nonzero scalar. Therefore, the corresponding estimate is defined only within a scalar factor. The algebraic least-squares estimator is derived from the cost function:

$$J_{\mathrm{ALS}}(\boldsymbol{\tau}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \|\boldsymbol{\tau}\|^{-2} \sum_{i=1} n\boldsymbol{\tau}^{\mathrm{T}} \boldsymbol{A}_i \boldsymbol{\tau} \tag{5}$$

where $\boldsymbol{A}_i = \Gamma(\boldsymbol{x}_i)\Gamma(\boldsymbol{x}_i)^{\mathrm{T}}$. Here, each $\boldsymbol{\tau}^{\mathrm{T}}\boldsymbol{A}_i\boldsymbol{\tau}$ is equal to the square of the algebraic distance $|\boldsymbol{\tau}^{\mathrm{T}}\Gamma(\boldsymbol{x}_i)|$. If the individual datum $\boldsymbol{x}_i$ is represented by (m_i, m_i'), the algebraic distance between the datum and a candidate fundamental matrix $\boldsymbol{F}$ can be written as $|m_i'\boldsymbol{F}m_i|$.

We let

$$\boldsymbol{U} = [\Gamma(\boldsymbol{x}_1), \ldots, \Gamma(\boldsymbol{x}_n)]$$

and the least-squares solution for $\boldsymbol{F}$ is the singular vector corresponding to the smallest singular value of the $\boldsymbol{U}$; that is, the last column of $\boldsymbol{V}$ in the SVD $\boldsymbol{U}_n = \boldsymbol{UDV}^{\mathrm{T}}$. The solution vector $\boldsymbol{F}^0$ found in this way minimizes $\|\boldsymbol{U}_n\boldsymbol{F}^0\|$ subject to the condition $\boldsymbol{F}^0 = 1$.

In general, the solution matrix $\boldsymbol{F}^0$ will not satisfy the singularity constraint of rank 2. Furthermore, the most convenient way to do this is to correct $\boldsymbol{F}^0$ using the SVD again. Let $\boldsymbol{F}^0 = \boldsymbol{UDV}^{\mathrm{T}}$ be the SVD of $\boldsymbol{F}^0$, where $\boldsymbol{D}$ is a diagonal matrix and $\boldsymbol{D} = \mathrm{diag}(r, s, t)^{\mathrm{T}}$ satisfying $r \geqslant s \geqslant t$. Then $\boldsymbol{F} = \boldsymbol{U}\mathrm{diag}(r, s, 0)\boldsymbol{V}^{\mathrm{T}}$ minimizes the Frobenius norm, which is the basic algebraic estimation of the fundamental matrix.

Generally, the solution above is the initial step of estimation, since it is only feasible for the 8-points and not optimal for the entries in the corresponding points set. Ref. [7] gives an iterative estimation based on the algebraic minimization algorithm. Its disadvantages are: 1) The times of iteration are unknown; 2) The estimation of the matrix is not optimal until the end of the process; 3) The approximation of the middle step result matrix in the iteration will not be uniformly convergent. However, in most reconstruction systems, the optimal estimation of the fundamental matrix is not a prerequisite. The major problem is how to find a good enough estimation. If we determine a norm error estimate to identify the approximation of optimization, we can develop algorithms to control the computation process automatically and find the proper estimation results.

3 Stripe Linear Constraint and HT Radius

Obviously, in a structured-light system, we can design some special patterns, and make the points in the projective plane have collinear constriction. Our scheme works as follows: Let $P_i (i = 0, \ldots, m)$ be points on the same stripe line in one camera plane and $P_i' (i = 0, \ldots, m)$ be their corresponding points on the projective plane, as shown in Fig. 3. Since P_i' is collinear in the projective plane, it satisfies

$$\boldsymbol{P}_i'^{\mathrm{T}} \boldsymbol{l} = \{x_i', \ y_i', \ 1\} \begin{Bmatrix} a \\ -1 \\ b \end{Bmatrix} = 0 \tag{6}$$

where $\{x_i', y_i, 1\}$ $(i = 0, \ldots, m)$ are the P_i's homogenous coordinates, and $\boldsymbol{l}$ is the linear coefficient vector. In the polar coordinates system, Eq. (6) can be rewritten as

$$\boldsymbol{P}_i'^{\mathrm{T}} \boldsymbol{l} = \{x_1', y_1', \ 1\} \begin{Bmatrix} \cos\theta \\ \sin\theta \\ \rho \end{Bmatrix} = 0 \tag{7}$$

In Eq. (7), point P_i' in the projective plane is transformed

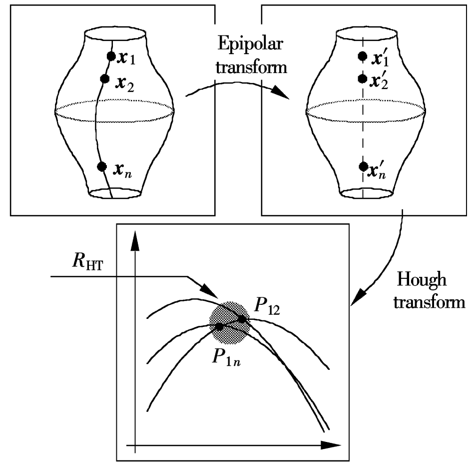

Fig. 3 Hough transform radius

into the corresponding sinusoids in the space of (θ, ρ), called the Hough transform. Since $P_i'(i = 0, \ldots, m)$ is collinear, the sinusoids must have one point of intersection in $[0, \pi]$. In fact, because the estimation result of the fundamental matrix is not optimal, the projective points will not be exactly collinear. In this case, the sinusoids will have several intersection points in a small range of $[0, \pi]$. The radius of the intersection area can be denoted as R_{HT}, called the HT radius. If the estimation matrix is more approximate to optimality, R_{HT} is closer to zero. Therefore, the HT radius is the function of the estimation result and a new fast criterion of optimization of the fundamental matrix estimation. The major steps of the algorithm to calculate the HT radius are as follows:

1) Let $\boldsymbol{F}$ be the estimation result matrix, and $\{P_{kl} \mid k = 1, \ldots, m\}$ be the points set on the stripe l in the image plane. Substituting $\boldsymbol{F}$ and $\{P_{kl}\}$ in Eq. (1), the corresponding points set $\{P_{kl}' \mid k = 1, \ldots, m\}$ on the projective planar can be obtained.

2) Using the Hough transform, the points set $\{P_{kl}' \mid k = 1, \ldots, m\}$ can be mapped to the Hough space, and the intersection set $\{T_{kl}\}$ in the range of $[0, \pi]$ on the plane (θ, ρ) of the sinusoid can be calculated. The cluster radius of T_{kl} is denoted as R_{HT}^l.

3) Repeat the above steps on other stripes in the image. The maximum value of R_{HT}^l is the R_{HT} of the estimation matrix $\boldsymbol{F}$.

Logically, the estimation optimization problem is converted to a cluster analysis problem, and R_{HT} is defined as the cost function for estimation.

$$J_{\mathrm{RHT}}(\boldsymbol{\tau}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \sum_{i=1}^{n} \|R_{\mathrm{HT}}(\boldsymbol{x})\|^2 \tag{8}$$

Following previous work, we present that the HT radius is the fast norm to check the optimal level of the current estimation, and can be used in the function automatically. The advantages of this method lie in: 1) The partial part of the corresponding points set, rather than the entire set, is calculated; 2) Each step of the process can be divided into several parallel computing parts in order to increase the efficiency of implementation in multi-core systems. The algorithm can be regarded as a variant of the Newton-Raphson method, shown as follows:

1) Find an initial approximation $\boldsymbol{F}^0$ for the fundamental matrix using the normalized 8-point algorithm.

2) Find the right null-vector $\boldsymbol{e}^0$ of $\boldsymbol{F}^0$.

3) Start with the estimate $\boldsymbol{e}^i \leftarrow \boldsymbol{e}^0$ for the epipolar, and compute the matrix $\boldsymbol{E}^i$.

4) Find the vector $\boldsymbol{f}^i = \boldsymbol{E}^i \boldsymbol{m}^i$ that minimizes Hough transform radius R^i_{HT}, which defines a mapping $\boldsymbol{e}^i \rightarrow R^i_{\mathrm{HT}}$.

5) Use the Levenberg-Marquardt algorithm to vary $\boldsymbol{e}^i$ iteratively so as to minimize $\| R^i_{\mathrm{HT}} \|$.

6) Do convergence until the desired fundamental matrix $\boldsymbol{F}$ is recovered.

4 Experiments

Our algorithms have been tested in a structured-light system with stripe code patterns. All experiments are done on a workstation with dual Intel Xeon 1 GHz processors with 1 GB of RAM and a Matrox Parhelia 128 MB PCI-Express graphics card. The projector is an Epson EMP500 with SVGA 2000 ANSI Lumen illumination. Captured through a CE-7861C CCD camera at a resolution of 352 × 288, the original color images are first converted to 8-bit gray-scale images before processing. The software system is developed with Borland C++ Builder 6 IDE on Windows XP.

Our experiments proceed as follows. A realistic stereo camera configuration is obtained when the camera and the projector's optical axes are not co-planar, and their intrinsic parameters are slightly different. 3D points are then projected onto the images to generate many pairs of corresponding points. A range of tests are then conducted; each is carried out with respect to an average level of noise a. For a given test, each image point is perturbed by adding zero-mean Gaussian noise of standard deviation a independently of each of the two coordinates. Each method is then supplied with these noise matching points and challenged to compute the fundamental matrix. For each a, the fundamental matrix is computed 50 times from a specific set of 96 corresponding points, with new perturbations being added each time. For each fundamental matrix obtained, an error measure is computed as the sum of the distances of the underlying true points and the epipolar lines derived from the estimated fundamental matrix, in both the left and right images. A composite error measure is then obtained by averaging this error over all 50 trials. This entire process is repeated for different average levels of noise (varying from 0.25 to 3 pixels in steps of 0.25). Fig. 4 shows the average epipolar-distance pixel errors obtained by each method. The tests reveal that the normal algebraic algorithm and the RHT algorithm perform identically. In our experiments, therefore, RHT succeeds in minimizing errors as well as the normal algorithm did.

Finally, an indicative reconstruction test is carried out. Corresponding points are extracted from images of an object in the structured-light system. The associated fundamental matrices are computed using the RHT algorithm. A self-calibration procedure is then used to determine the relative orientation of the cameras, with the camera intrinsic parameters having been precalibrated in the laboratory. Fig. 5 shows the reconstruction results of a tiger mask. Fig. 5(a) is the original frame with stripe codes. Fig. 5(b) is the point cloud of the reconstruction results from the same perspective.

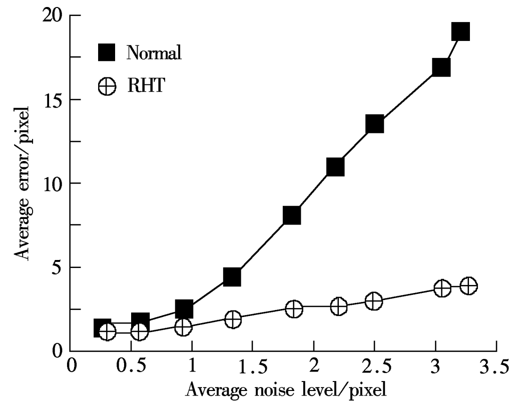

Fig. 4 Fundamental matrix estimation error vs. average noise level

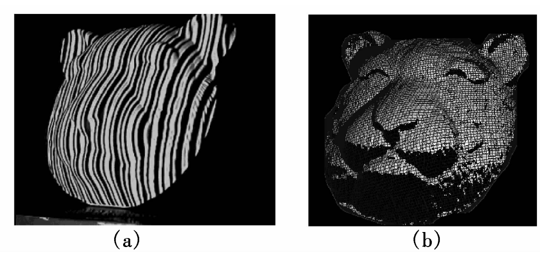

Fig. 5 Reconstruction of a tiger mask. (a) A frame of captured images of a tiger mask; (b) The point cloud of the reconstruction results

5 Conclusions

The significant contributions obtained from this paper are enumerated as follows:

1) The HT radius is introduced as a new fast criterion of optimization of the fundamental matrix estimation for converting the estimation optimization problem to a cluster analysis problem in a Hough transform space.

2) The new algorithm based on stripe constraints is presented. The algorithm is an effective variant of the traditional calculation of the fundamental matrix. The practicality and performance of the algorithms are verified by experiments in this paper.

References

[1] Mittal A, Davis L S. A general method for sensor planning in multi-sensor systems: extension to random occlusion [J]. *International Journal of Computer Vision*, 2008, **76**(1): 31 − 52.

[2] Tang C Y, Chou H L, Wu Y L, et al. Robust fundamental matrix estimation using coplanar constraints [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2008, **22**(4): 783 − 805.

[3] Hua M, McMenemy K, Ferguson S, et al. Epipolar geometry estimation based on evolutionary agents[J]. *Pattern Recognition*, 2008, **41**(2): 575 − 591.

[4] Huang J F, Lai S H, Cheng C M. Robust fundamental matrix estimation with accurate outlier detection[J]. *Journal of Information Science and Engineering*, 2007, **23**(4): 1213 − 1225.

[5] Chojnacki W, Brooks M J. On the consistency of the normalized eight-point algorithm [J]. *Journal of Mathematical Imaging and Vision*, 2007, **28**(1): 19 − 27.

[6] Brandt S. On the probabilistic epipolar geometry [J]. *Image*

and Vision Computing, 2008, **26**(3): 405 − 414.

[7] Tang C Y, Wu Y L, Lai Y H. Fundamental matrix estimation using evolutionary algorithms with multi-objective functions [J]. *Journal of Information Science and Engineering*, 2008, **24**(3): 785 − 800.

基于条纹边界约束的基本矩阵快速估算

杨安康[1]　刘京南[1]　内田敬久[2]　内田悦行[2]

([1] 东南大学自动化学院,南京 210096)

([2] 爱知工业大学情报通讯工学科,日本爱知丰田 470-0392)

摘要:为了提高计算机视觉领域中核心问题之一的基础矩阵估算的效率,基于条纹边界编码约束实现了一种快速估算方法. 与传统的基于代数最小二乘法算法不同,该算法利用 Hough 变换将 Hough 半径作为最优化过程的最小化因子. 在特定条纹边界编码的结构光投影系统模型下,利用条纹编码的共面性构造线性约束,采用 Hough 变换将同码像素映射到 Hough 空间,其交点的半径可作为最优化评价函数的最小化因子. 再通过 Levenberg-Marquardt 最优化迭代过程估算出基本矩阵的全局最优解. 实验结果表明了该算法的正确性,并证明了其可有效提高估算精度与效率.

关键词:基本矩阵;结构光;条纹编码模式;双目视觉

中图分类号:TP391