MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan
8-6

Feature Point Matching of Stereo Images by using a Projective Invariant

Yusuke Ishida
Meijo University
Nagoya, Aichi, Japan
m0432001@ccmailg.meijo-u.ac.jp

Toshimitsu Tanaka
Meijo University
Nagoya, Aichi, Japan
toshitnk@ccmfs.meijo-u.ac.jp

Noboru Sugie
Meijo University
Nagoya, Aichi, Japan
sugie@ccmfs.meijo-u.ac.jp

Yuji Sagawa
Meijo University
Nagoya, Aichi, Japan
sagawa@ccmfs.meijo-u.ac.jp

Abstract

In Image-Based Modeling, which reconstructs 3-dimensional structures and extracts a texture from 2-dimensional images, one of the greatest problems is to match feature points of two images. The well-known template matching is used for automatically relating individual points in different images. However, since it matches each image in which a viewpoint position differs, incorrect feature points are often related. In this study, we propose the automatic matching technique of feature points that is not sensitive to parallax.

We aim to improve the accuracy of feature point matching by using a projective invariant on a plane surface. In order to calculate the projective invariant, we use collinear four points or coplanar five points. Firstly, in each image corners of objects are picked out as feature points. Next, an edge image is created from the original image. The connective relation of edge points between feature points is measured. All of the collinear four feature points and the coplanar five feature points are estimated from the connective relation, and a projective invariant is calculated for each set of points. Finally, all projective invariants of one image are compared to those of another image and feature points with an equal projective invariant are associated. Then we confirm the effectiveness of the proposed technique by the experiments.

1 Introduction

For modeling realistic shapes and textures manually, a certain amount of knowledge and experience are required. Moreover, much time and labor is needed. To reduce the modeling cost, Image-Based Modeling was developed. This technique reconstructs 3-dimensional structures and extracts textures from photographs using the stereo vision method [1][2][3]. The stereo method reconstructs 3-dimensional structures by corresponding feature points of one image to the points of another image. The template matching is a popular method for automatically corresponding feature points. However, if viewpoint positions of two images much differ, incorrect matching often occurs. Thus, it is not easy to perfectly and automatically match feature points of stereo images.

In this study, we propose the automatic matching technique of feature points that is not sensitive to parallax. We aim to raise the accuracy of feature point matching by using a projective invariant, which is a steady value even if the viewpoint changes [4][5]. Although there are various kinds of projective invariants, we use projective invariants for four points on a line and that for five points on a plane surface.

The flow of the whole system is shown in Fig.1. First, feature points are detected by Harris operator [6][7]. Next, the ratio of the edge points between feature points is measured, and the connective relation of the feature points is determined. Then, collinear four points and coplanar five points are estimated by referring the connection relation. Finally, feature points are matched by comparing projective invariants, which are calculated in each image.

If number of coplanar feature points is less than 5, they are not related. However, their points are matched later. Our method can accurately relate feature points even if two images have big parallax. Thus, correct fundamental matrix of perspective projection can be computed without estimation methods like RANSAC. By using this matrix, the remained points can be related efficiently.

We are planning to apply our method to indoor scenes. Since many artificial objects that consist of straight lines and flat surfaces exist in the scenes, sufficient number of matching points can be obtained.

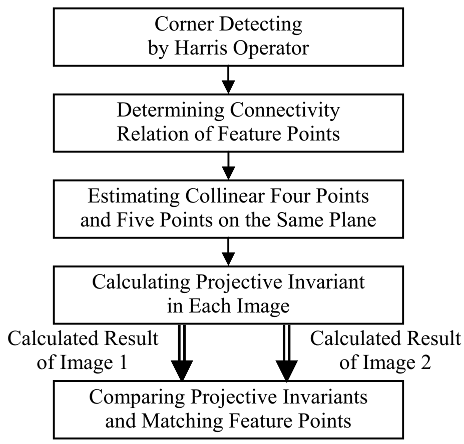

Figure 1. Flow of the whole system

2 Projective Invariant on a Plane

2.1 Projective Invariant of Collinear four Points

Let $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and $\boldsymbol{x}_4$ be any four distinct collinear points on an image plane. The cross ratio

$$\frac{L(\boldsymbol{x}_2,\boldsymbol{x}_3)L(\boldsymbol{x}_1,\boldsymbol{x}_4)}{L(\boldsymbol{x}_1,\boldsymbol{x}_3)L(\boldsymbol{x}_2,\boldsymbol{x}_4)} \qquad \text{(Eq.1)}$$

is considered, where $L(\boldsymbol{a}, \boldsymbol{b})$ is the distance between point $\boldsymbol{a}$ and $\boldsymbol{b}$. Such a cross ratio calculated from collinear four points constitutes a projective invariant. As shown in Fig.2, collinear four points $\boldsymbol{p}_i$ (i=1,2,3,4) in 3-dimensional space are projected onto the distinct projective plane π and π'. The point $\boldsymbol{x}_i$ is the projection of $\boldsymbol{p}_i$ on the plane π, and the point $\boldsymbol{x}'_i$ is the projection of $\boldsymbol{p}_i$ on the plane π', respectively. The cross ratio calculated from $\boldsymbol{x}_i$ is equal to the cross ratio calculated from $\boldsymbol{x}'_i$.

The possibility that values of Eq.1 computed from different sets of points are same is very low. Thus, if projective invariants calculated from collinear four points on different images are equal, the points are the projections of the same points in 3-dimensional space. Accordingly such points are related between the two images.

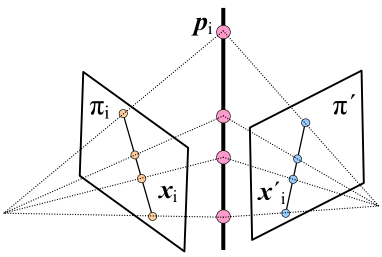

Figure 2. Collinear four points and projective planes

2.2 Projective Invariant of Non-collinear Points on a Plane

Let $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, $\boldsymbol{x}_4$ and $\boldsymbol{x}_5$ be any five distinct points that are non-collinear on a plane. The expression

$$\frac{\det([\tilde{\boldsymbol{x}}_1 \quad \tilde{\boldsymbol{x}}_2 \quad \tilde{\boldsymbol{x}}_4])\det([\tilde{\boldsymbol{x}}_1 \quad \tilde{\boldsymbol{x}}_3 \quad \tilde{\boldsymbol{x}}_5])}{\det([\tilde{\boldsymbol{x}}_1 \quad \tilde{\boldsymbol{x}}_2 \quad \tilde{\boldsymbol{x}}_3])\det([\tilde{\boldsymbol{x}}_1 \quad \tilde{\boldsymbol{x}}_4 \quad \tilde{\boldsymbol{x}}_5])} \qquad \text{(Eq.2)}$$

is considered, where $\tilde{\boldsymbol{x}} = [x \quad y \quad 1]^{\mathrm{T}}$, and the coordinate of $\boldsymbol{x}$ is (x, y). $\det([\tilde{\boldsymbol{a}} \quad \tilde{\boldsymbol{b}} \quad \tilde{\boldsymbol{c}}])$ is area of the triangle which points $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ as vertex make. The value of this expression also constitutes a projective invariant.

If object shapes are simple, the number of the case that four feature points are on a line is not so many. In order to increase the number of points matched between two images, we use the projective invariant calculated from five points on a plane.

3 Detecting Feature Points

3.1 Harris Operator

It is necessary to detect feature points from each image first. We use Harris operator that detects the corner points of objects in an image by comparing luminosity of pixels. Let $f(x, y)$ be the luminosity value of an image at a point (x, y), the differentiation values

$$D_x = \frac{\partial}{\partial x} f(x,y), \quad D_y = \frac{\partial}{\partial y} f(x,y)$$

are calculated by Sobel operator of x direction and y direction, respectively. Using the values, matrix $\mathbf{M}$ is computed by the following equations.

$$A = \sum D_x^{\,2}, \quad B = \sum D_y^{\,2}, \quad C = \sum D_x D_y$$

$$\mathbf{M} = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$$

Let λ_1 and λ_2 be eigen values of matrix $\mathbf{M}$, the following expression is computed.

$$H = \lambda_1\lambda_2 - \alpha(\lambda_1 + \lambda_2)^2$$

Here α is a constant which means the accuracy of detection. If H is larger than the threshold value, the point will be detected as a corner.

3.2 Measurement of the ratio of edge

In preparation for searching collinear four points and coplanar five points, we determine the connectivity of feature points. We evaluate the connection of two feature points using the ratio of the edge pixels between them. The number of edge pixels on the line segment connected the feature points are counted up. Then ratio of the edge pixels is computed by (the number of the edge points on a line segment) / (the number of pixels on a line segment). If the value is larger than the threshold, we judged that the two points are connected. By integrating the judgments, the connectivity of any feature points can be understood as a kind of graph.

3.3 Detection of Collinear Points

In order to search collinear four points, we track feature points. The connectivity of the feature points defined above is described with the adjacency matrix. The algorithm for tracking the feature points is shown below.

1. *Detected feature points are sorted in order of their X positions (from the left to the right).*

Feature points are tracked from left to right. The next feature point can be efficiently chosen by the sorting.

2. *The feature point is tracked by a recursive call.*

The tracking is recursively called. The flow of the process is illustrated in Fig. 3. First, we start the tracking from the most left feature point which is recorded in Fig. 3(A) as 1. Next, we move to the point 2 that is the nearest feature point connected to the point 1, by the recursive call (B). We track from the feature point 2 to the 3 similarly. In addition, the angle that line segments 1-2 and line segments 2-3 make is computed, and it is judged whether the new point is on the connecting line of points 1 and 2 or not (C). When it is not on the straight line, another point is searched. If it is on the straight line, the next feature point is tracked similarly (D). Note that the feature point 2 has branch. The effect of a recursive call is shown here. That is, by the track back, we return to the feature point 2 and can search for the next point (E). Although we are going to track the feature point 5, we don't track that point actually, since it is judged non-straight line by the angle check. Thus we do not go to (F). The tracking process ends here.

3. *If four or more points are tracked, we output the path.*

The tracked points are output as a set of collinear points. If more than four points are tracked, all combinations of four points are generated, and a projective invariant is calculated for each combination.

3.4 Detection of Points on a Plane Surface

In order to select five points on a plane, we propose two methods. The first method tracks five feature points which are not on a line. If the path connecting the five points is closed such that the points make a pentagon, the points are selected because they are usually on a plane. The second method finds a quadrangle in the same way, because many indoor objects, such as doors, window frames, desks, bookcases, and etc., consist of quadrangles. And four vertices of the quadrangle and one point in its inside are selected. The algorithm for detecting a pentagon or a quadrangle is almost same to the algorithm for detecting collinear four points. The algorithm is shown below.

1. *The feature point is tracked by a recursive call.*

The feature point is tracked by a recursive call like the method for collinear four points. However, in this case points that are not on a straight line are tracked. For example, the processes (A) and (B) in Fig. 3 are done similarly. However the feature point 3 is not traced because the angle of the line segment 1-2 and the line segment 2-3 is almost 180 degree in (C). Therefore, feature point 5 is tracked as shown in (F).

2. *If four or five points are tracked and they constitute closed path, we output the path.*

The closed path including four or five points is output because the points are usually on the same plane. If closed four points are tracked, the fifth point is selected in the inside of the quadrangle. All possible combinations are generated and projective invariants are computed for each combination.

4 Matching feature points

The projective invariants are computed and compared independently in the following three cases; (1) collinear four points, (2) vertices of a pentagon, (3) vertices of a quadrangle and a point inside the quadrangle. For each set of feature points a projective invariant is computed by using Eq.1 in the case (1) and by using Eq.2 in (2) and (3). Then projective invariants of an image are compared to those of another image. If a same value appears in both images, the feature points used to compute the value in each image are related. Strictly speaking, values of the projective invariants in two images are not equal because of the quantization error of positions. Thus if difference of the two projective invariants is less than the threshold, it is considered that they are equal.

5 Experiments

In order to test the proposed method, we applied it to a computer graphics image. The results are shown in Fig. 4. The object is a plane in 3-dimensional space on which the texture is mapped. The images (a) and (b) in Fig. 4 were created by projecting the plane from different viewpoints.

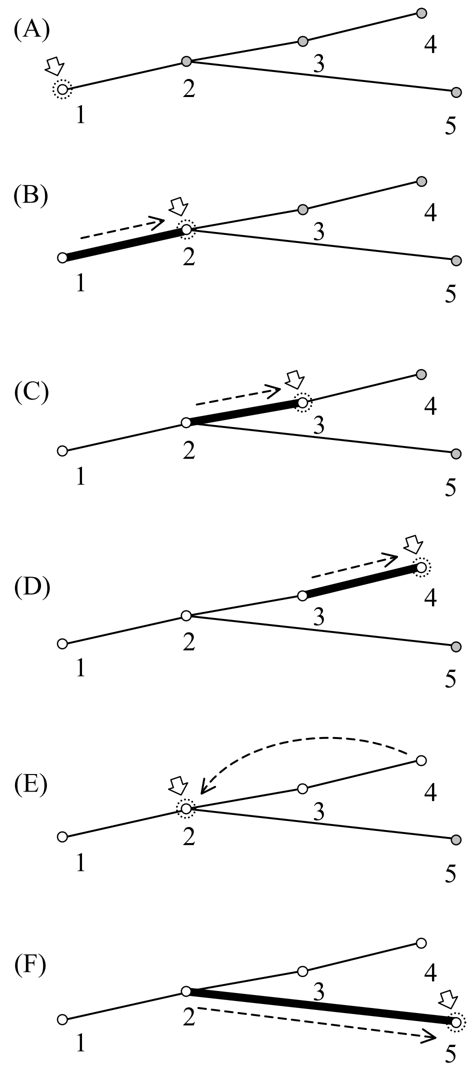

Figure 3. The flow of tracking feature points

The viewpoint of (b) is rotated 54 degrees from (a). The image resolution is 800×600 pixels. In the images, the small squares show the feature points detected by Harris operator. The feature points indicated with the same number in images (a) and (b) are related. All feature points were correctly related although camera positions of the images are much different.

The results of photographs are shown in Fig. 5. The two photographs show the same LCD monitor from the different viewpoints. On the monitor, the “e” mark is displayed. The image resolution is 1000×750 pixels.

As a result of the feature point matching, almost points were correctly paired in two images. However, points 4 and 7 were wrongly related. This cause is that the difference of projective invariants of a correct pair and wrong pair is very close value. This difference is 0.00095.

The projective invariants were computed from the points 0, 1, 8, 9 and 4 and points 0, 1, 8, 9 and 7, respectively. However since the points 0, 1, 8, and 9 are corners of the monitor frame, the plane defined by the points are a few millimeters forward on the LC screen surface. Thus, in the strict sense, neither point 4 nor point 7 are on the plane. This probably caused the counterturn of projective invariants.

Since our system selects smallest pair and relates its feature points, the wrong matching occured. This problem may be solved by using image information around a feature point. For example, after two feature points are matched, colors and textures around the points are compared to confirm the matching. If they are vastly different, the next matching is selected.

(a)

(b)

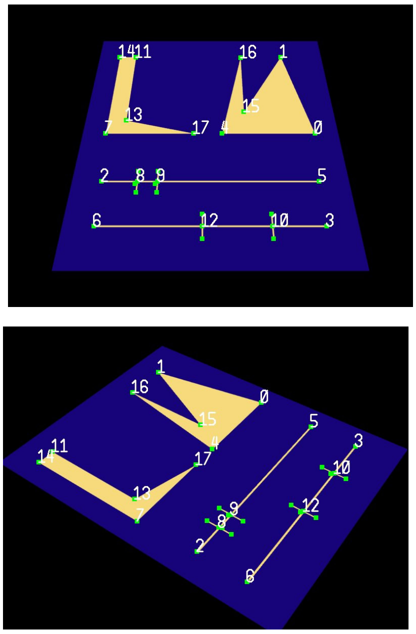

Figure 4. The result of the computer graphics image

(a)

(b)

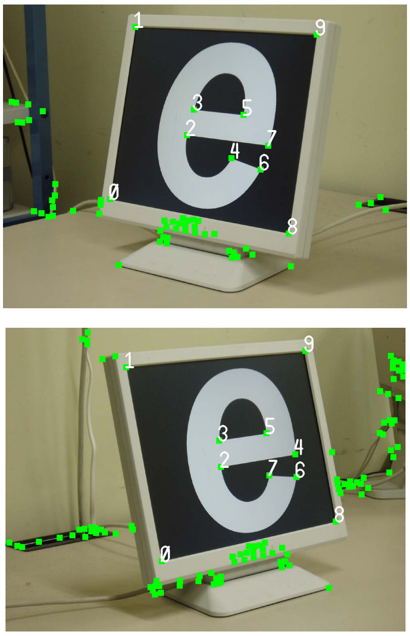

Figure 5. The result of the photograph image

6 Conclusion

In this paper, we proposed the method that matches feature points between two images by using projective invariant. The method is not sensitive to parallax. In order to detect feature points, we use Harris operator. Then, we use the projective invariant on a plane, which requires collinear four points or coplanar five points to calculate. The method to find these points on an image was described.

The method determines connective relation between the feature points first. And then it tracks the connective relation in order to detect feature points that satisfy the condition. If the points are detected, the projective invariants are computed and compared in two images. If the values are equal, the points are matched.

The effectiveness of the proposed method was confirmed in the experiments. All feature points in CG images were correctly related. Almost all points in photographs were also correctly matched; however, a few points were missed.

7 Future works

The proposed technique relates collinear four feature points existing in both of two images. It also relates coplanar five points. However, it cannot relate feature points not on a plane surface. The remained points are matched by using the epipolar geometry after the fundamental matrix of the perspective conversion is estimated from the related feature points.

The fundamental matrix can be calculated from intrinsic parameters and extrinsic parameters of a camera. The intrinsic parameters are known from spec of the camera. The extrinsic parameters can be obtained from corresponded feature points with high reliability. We think that the reliability can be determined by difference of projective invariants, distance between feature points, area of quadrangle and pentagon, or color information of texture.

We have to improve this technique more robustly, since the results of photograph images are not perfect. Finally, we will develop the method reconstructing 3-dimensional structures automatically.

References

[1] G.Xu, 3D CG from Photographs: Image-Based Modeling and Rendering, Kindai Kagaku Sha, January, 2001. (Japanese)

[2] G.Xu and S.Tujii, Three-Dimensional Vision, Kyoritsu Shuppan, April, 1998. (Japanese)

[3] L.McMillan and S.Gortler, Image-Based Rendering. In Computer Graphics Quarterly, ACM SIGGRAPH Vol.33, No.4, November 1999. http://www.siggraph.org/publications/newsletter/v33n4/contributions/mcmillan.html

[4] J.Sato, Computer Vision – Geometry of Vision –, Corona publishing, May, 1999. (Japanese)

[5] E.B.Barrett and P.M.Payton, General Methods for Determining Projective Invariants in Imagery, CVIP, Vol. 53, No. 1, January, pp. 46-65, 1991.

[6] C.Harris and M.Stephens, A Combined Corner and Edge Detector, In Proc. 4th Alvey Vision Conf., Manchester, pp. 147-151. 1988.

[7] D.Csetverikov, Basic Algorithms for Digital Image Analysis: a course, http://ssip2003.info.uvt.ro/lectures/chetverikov/shape_analysis.pdf

[8] E.Vincent and R.Laganiere, Matching Feature Points in Stereo Pairs: A Comparative Study of Some Matching Strategies, In Machine Graphics & Vision, vol. 10, no. 3, pp. 237-259, 2001.

[9] D.A.Forsyth and J.Ponce, Computer Vision A Modern Approach, Person Education, 2003.