Journal of Southeast University (English Edition) Vol.19 No.3 Sept. 2003 ISSN 1003—7985

A method of automatic plane detection without random search

Li Zhongke Yang Xiaohui Wu Lenan

(Department of Radio Engineering, Southeast University, Nanjing 210096, China)

Abstract: Plane detection is a prerequisite for many computer vision tasks. This paper proposes a new method which can automatically detect planes from two projective images. Firstly, we modify Scott's feature point matching method by post-processing its result with the concept of similarity, and then get the lines matching according to feature points matching based on the approximate invariance of the features' distribution between two images. Finally, we group all feature points into subsets in terms of their geometric relations with feature lines as initial sets to estimate homography rather than by a random search strategy (like RANSAC) as in most existing methods. The proposed method is especially suitable to detecting planes in man-made scenes. This method is validated on real images.

Key words: plane detection; feature matching; plane homography; computer vision

Planes are commonly used in various vision tasks because of their attractive geometric properties and their familiarity in man-made environments. They have been successfully employed in many fields such as feature matching and grouping, camera self-calibration, obstacle estimation, 3-D reconstruction and scene analysis, object recognition, robot visual navigation and so on[1]. Despite their popularity, most vision algorithms based on the existence of planes rely on manual or even unspecified schemes for detecting such planes in images. For instance, assuming the two homography matrices of two planes in space are available, G. Xu, et al.[2] proposed a linear algorithm for recovery of plane equations and motion between the two cameras simultaneously. Plane detection is a prerequisite for these algorithms.

The state-of-art methods for plane detection are based on the extraction and matching of sparse geometric features from images[1,3-6]. Some firstly get the feature matching automatically or manually, and then estimate the plane homography matrix[1,3,4]; others match the feature and estimate the homography at the same time[6,7]. However, most of such algorithms employ random sample consensus (RANSAC)[3,4,6,7] or other like methods[1,5]. In fact, RANSAC performs a tentative global searching strategy, which proceeds by repeatedly estimating solutions from a minimal set of points gathered from the data, and by testing the complete set of data to find a support for each solution, the solution that maximizes the number of the support is selected[3,8]. But the minimal set is completely randomly chosen without the help of any knowledge. Plenty of repetition is required to get the exact solution consistently. Actually in most man-made environments, there are line-edges between the joined planes. This paper employs this property to assist in plane detection. In terms of the geometric relation between the matched points and lines between two images, we group the feature points into several subsets as initial sets to estimate the plane homography, so in order to avoid the cost of global searching.

The contributions in this paper are that a new method is proposed to get the line correspondence by the matched feature points and an initial set is given to estimate plane homography by grouping feature points according to the geometric relations between feature points and feature lines.

1 Feature Extraction

1.1 Feature point extraction

Firstly we extract feature points based on window auto-correlation corner detection. The center of the window is defined as the exact location of the feature point. If a window area can be tracked and matched reliably, the gradient of the image in this window should be great enough in both two independent directions[9]. The window area conformed to the requirement above can be derived as follows:

$$\lambda_{\min}(M) > \tau, M = \int_{W(x_0, y_0)} \nabla I \nabla I^{\mathrm{T}} \mathrm{d}x \mathrm{d}y \tag{1}$$

Received 2003-04-11.

Foundation item: The Jiangsu Key Lab of Image Processing and Image Communication (K02091).

Biographies: Li Zhongke (1976—), male, graduate; Wu Lenan (corresponding author), male, doctor, professor, wuln@seu.edu.cn.

where λ_{min} denotes the smallest eigenvalue; W is a Gauss smoothing window; I and ∇I denote image intensity and image intensity gradient, respectively; that is, the smallest eigenvalue of M must be larger than some defined threshold. In this paper, based on the idea above, the Plessey corner detector[10] is employed, the measurement of which is defined by

$$\Delta = \alpha\beta - K(\alpha - \beta)^2 \tag{2}$$

where α and β are two eigenvalues of M ; K is a constant to balance the detection and localization performance of the corner detector. For two images to be tackled, Eq.(2) is computed at every point in each image. The locations, where measurements are above a threshold, are chosen as the candidates of feature points. In order to avoid the latter mismatching due to the immoderate denseness of feature points, if two points are nearer than a small distance, we retain only one with the larger measurement.

1.2 Feature line extraction

The Hough transform is adopted to extract feature lines in this paper. First, the image edge is obtained using the Canny detector. For edge image, the Radon transform is computed for each angle θ:

$$R_\theta(x') = \int_{-\infty}^{\infty} f(x'\cos\theta - y'\sin\theta, x'\sin\theta + y'\cos\theta)\mathrm{d}y' \quad \theta \in [0,2\pi) \tag{3}$$

where x'-y' is the coordinate axis after x-y is rotated θ courter-clockwise. After obtaining the accumulator function $R(\theta, x')$, we get the local maximum of $R(\theta, x')$ as the candidates of the lines. If two candidates are too near, the operation is the same as that performed on the features points. Considering one line in the image has two projections in the range $[0, 2\pi)$, we retain only one with a slope less than π.

2 Feature Matching

2.1 Feature point matching

In this paper, the method for associating two patterns[11] is imported to solve the problem of feature points matching. For the case of real image pairs, some of the matched points are nonsensical via this method. The concept of similarity is adopted from Ref.[11], but it requires computing $m \times n$ correlations. The number of correlations is less than m when we post-process the result of feature points matching. Experiments show that the performance is comparative to that in Ref.[11]. Alternatively, other feature points matching strategies can also be adopted instead.

2.2 Line matching

Based on the one-to-one correspondence of feature points between two images, we propose a new method for line matching between two images:

First of all, we define a direction function $L(P)$ (see Fig.1). P is the point (P_x, P_y), and L denotes a line function $L(x, y)$. Then we define another function named point-to-line function $\delta(P, L)$:

$$\delta(P, L) = \mathrm{sgn}(L(P)) = \begin{cases} 1 & L(P) > 0 \\ 0 & L(P) = 0 \\ -1 & L(P) < 0 \end{cases} \tag{4}$$

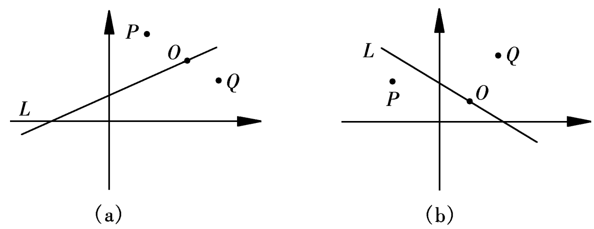

Fig.1 Direction function: $L(P) > 0$, $L(O) = 0$, $L(Q) < 0$. (a) $0 \leqslant \theta \leqslant 90°$; (b) $90° < \theta < 180°$

If two lines from two images are counterparts, the distribution of the matched feature points should be almost the same at each side of these lines. So we define the correspondence of line i in the second line set as follows:

$$C(i) = \begin{cases} \operatorname{argmax}_j\left(\sum_{k=1}^{K} \delta(P_{1k}, l_{1i})\delta(P_{2k}, l_{2j})\right) & \kappa > \lambda K \\ -1 & \kappa < \lambda K \end{cases} \tag{5}$$

where $\kappa = \max_j\left(\sum_{k=1}^{K} \delta(P_{1k}, l_{1i})\delta(P_{2k}, l_{2j})\right)$; λ is a constant less than 1 and comparative to 1. -1 denotes that the line i has no correspondence in the second line set. The hard decision in $C(i)$ is to tackle the change of the point-to-line relations due to feature points belonging to the different planes. If the scene has only one plane, the hard decision can be eliminated.

3 Plane Detection

Assume that two images are obtained under projective projection. It is widely known that the projective transformation of a plane between two images can be described by a homography matrix $\boldsymbol{H}$. More specifically, if p is the projection of a point in one view on the plane and p' is the corresponding projection in the second view, then the two projections are related by a linear projective transformation:

$$p' \cong \boldsymbol{H}p \tag{6}$$

where "$\cong$" denotes equality up to a scale.

In general, a plane has straight-line-edges in man-made environments. In this paper, we employ these edges to group feature points. The feature lines extracted from only one image are often in a clutter and don't always have practical meanings. It is why we import the line correspondence mentioned in section 2. After line correspondence, the sum of the number of lines is decreased and the geometric meanings of the lines are clearer than before. So we can group feature points in one image in terms of the point-to-line relation as follows:

1) Code each of all the matched N points in terms of the point-to-line relation function for all K lines. If the relation function between the points and a line is greater than 0 or equal to 0, set the corresponding bit to 1, otherwise set to 0, and get a code of K bits for each feature point (altogether N codes).

2) Group the points with the same code into one point set. We invalidate the point sets whose numbers are less than the minimum required. And then take the valid point sets as the initial sets to estimate the plane homography.

We first estimate $\boldsymbol{H}$ using each valid initial set and compute the distance $|p' - \boldsymbol{H}p|$ for all feature points. The feature points of which the distances are less than a threshold are regarded as the coplanar points, from which we estimate the $\boldsymbol{H}$ repeatedly until the member of the coplanar points is invariable. Then we choose the set, which has the most members among all the coplanar point sets, as the dominant plane and save it. We repeat the process above for the residuals of feature points.

4 Experimental Results

The performance of the proposed method has been evaluated on a set of test images. Experiments show that this method can detect the planes effectively and quickly after feature matching. In general, the iterative number of the homography estimation is less than 4 before the coplanar point sets become constant. Three representative results are given in this section.

Fig.2 illustrates the process and the result of plane detection for the image pairs-the first frame and the thirtieth frame from hotel sequence (image courtesy of CMU). The first figure is the extraction result of feature (including points and line) for the first frame. "+" represents feature points, and the white straight line means feature line (see Fig.2). The second figure and the third figure are the result of feature matching (the yellow numbers are the labels of feature points.

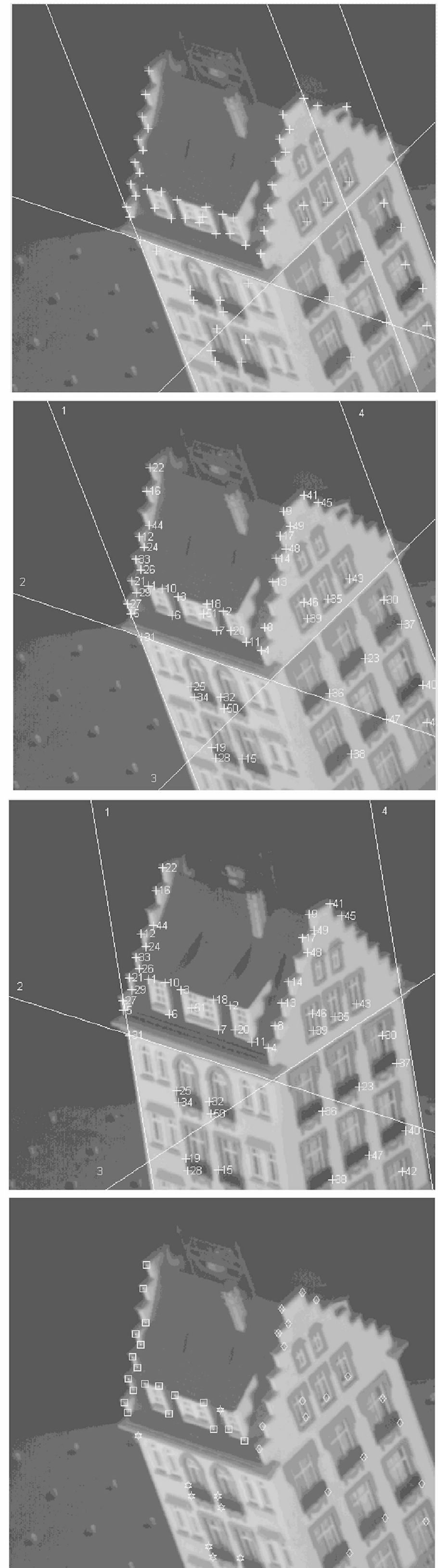

Fig.2 The process and the result of plane detection for the image pairs — the first frame and the thirtieth frame from hotel sequence (image courtesy of CMU)

The blue numbers are the labels of the white feature lines). The forth figure is the distribution of coplanar points (take the first frame as the example). The result of plane detection is that three planes are detected — the right wall, the roof and the left wall. Diamonds/squares/hexagons denote the first/second/ third planes, respectively, in the order of time when the planes are detected. Only one feature point labeled "2" is misgrouped.

Fig.3 illustrates the result of plane detection for the image pairs "Qiangongyuan". Left (benchmark) and middle are the result of feature matching (the signs have the same meanings as in Fig.2). Right is the distribution of coplanar points (for benchmark). The result of plane detection is that two planes are detected — the wall with two windows and the floor. Dimonds/squres denote the first/second, respectively, in the order of time when the planes are detected. In the end, the feature points labeled 1, 2 and 14 are left. No point is misgrouped.

Fig.4 illustrates the consequence of plane detection for the image pairs "Kejiguan". Left (benchmark) and middle are the result of feature matching (the signs have the same meanings as in Fig. 2). Right is the distribution of coplanar points (for benchmark). The result of plane detection is that three planes are detected (as shown in Fig.4). Diamonds squares/hexagons denote the first/second/third planes, respectively, in the order of time when the planes are detected. In the end, the feature points labeled 13, 19, 20 and 27 are left. No point is misgrouped.

Fig.3 The result of plane detection for the image pairs — "Qiangongyuan"

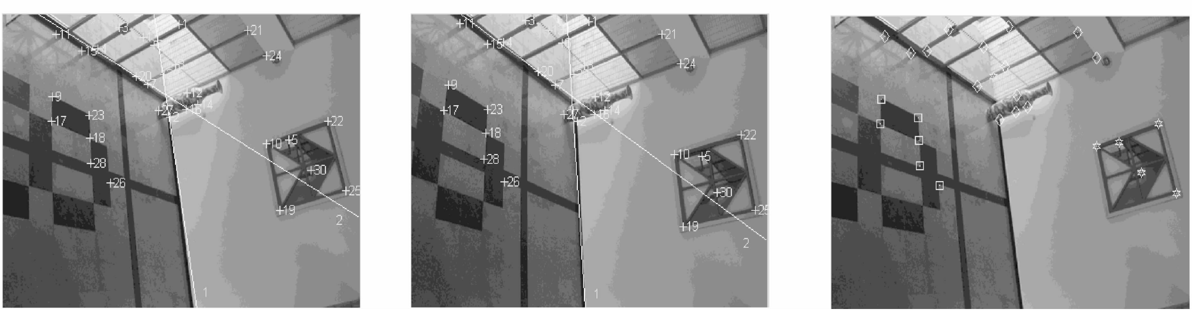

Fig.4 The result of plane detection for the image pairs — "Kejiguan"

5 Conclusion

In this paper we firstly match feature points and lines between two images, and then according to point-to-line geometric relations we group the feature points into subsets as the initial sets to estimate the plane homography. The experimental results show that the proposed method can avoid random searching effectively, detect the plane swiftly, and segment all the feature points into multi-coplanar groups. The proposed method is especially suitable to detect planes in man-made scenes.

References

[1] Lourakis M I A, Argyrosand A A, Orphanoudakis S C. Detecting planes in an uncalibrated image pair [A]. In: *BMVC* [C]. 2002, **2**: 587 - 596.

[2] Xu G, Terai J, Shum H Y. A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images [A]. In: *IEEE CVPR* [C]. 2000, **2**: 474 - 479.

[3] Bartoli A. Piecewise planar segmentation for automatic scene modeling [A]. In: *IEEE CVPR*[C]. 2001, **2**: 283 - 289.

[4] Bartoli A, Sturm P, Horaud R. *A projective framework for structure and motion recovery from two views of a piecewise planar scene*[R]. INRIA Research Report RR-4070. 2000.

[5] Pears N, Liang B. Ground plane segmentation for mobile robot visual navigation [A]. In: *IEEE/RSJ International Conf on Intelligent Robots and Systems*[C]. 2001. 1513-1518.

[6] Fornland P. Dominant plane detection for uncalibrated binocular vision [EB/OL]. http://www.nada.kth.se/cvap/ongoing-991214.html. 1998.

[7] Lourakis M I A, Halkidis S T, Orphanoudakis S C. Matching disparate views of planar surfaces using projective invariants [J]. *IVC*, 2000, **18**(9): 673-683.

[8] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Communications of ACM*, 1981, **24**(6): 381-395.

[9] Shi J, Tomasi C. Good features to track [A]. In: *IEEE CVPR*[C]. 1994. 593-600.

[10] Zheng Z Q, Wang H, Teoh E K. Analysis of gray level corner detection [J]. *Pattern Recognition Letter*, 1999, **20**(2): 149-162.

[11] Pilu M. Uncalibrated stereo correspondence by singular value decomposition [R]. HP Technical Reports, HPL-97-96. 1997.

一种无需随机搜索策略的自动平面检测方法

李中科　杨晓辉　吴乐南

（东南大学无线电工程系，南京 210096）

摘　要　本文提出一种新的能快速从2幅投影图像中自动检测平面的方法.首先通过引入互相关对Scatt特征点匹配结果作后处理得到可靠的点点匹配，然后根据2幅图片间特征分布的近似不变性从特征点匹配得到特征线匹配，继而将特征点按照与特征线的几何关系进行分组，作为初始集合估计平面单应矩阵，将特征点按所属平面分组.此方法能有效地避免随机的全局搜索.此方法尤其适于人造场景的平面检测问题.实验表明了此方法的有效性.

关键词　平面检测；特征匹配；平面单应矩阵；计算机视觉

中图分类号　TP391.4