A Generalized Computation Model for Plane Normal Reocery by Searching on Gaussian Hemisphere

Zhaozheng Hu [1, 2]
[1] Graduate School of Informatics
Kyoto University
Kyoto 606-8501, Japan
E-mail:zhaozheng.hu@gmail.com

Takashi Matsuyama [1]
[2] College of Information Science and Technology
Dalian Maritime University
Dalian 116026, China
E-mail: tm@vision.kuee.kyoto-u.ac.jp

***Abstract*—Computation of plane normal is crucial for single view based vision computation, such as metrology, pose estimation, localization, motion analysis, etc. While most of existing algorithms are based some specific scene constraints, e.g., parallel lines, control points or lines, specific shape, texture, etc., it is difficult to combine all possible geometric constraints into one framework. This paper presents a novel generalized computation model to recover plane normal by incorporating different geometric constraints. In this paper, we formulate the problem of plane normal computation from a set of geometric constraints as to maximize the likelihood that these constraints are satisfied. The likelihood that each constraint is satisfied is modeled by using the Gaussian function to follow four basic rules that we define. Computing of the maximum likelihood is accomplished by searching on the Gaussian hemisphere. The proposed algorithm has been tested with both simulation data and real image data. The experimental results from both simulation and real image data show that the proposed algorithm is accurate and reliable, which demonstrate that the proposed computation model is very practical and useful for many single view based vision computation applications.**

Keywords- Plane normal recovery, generalized computation model, Gaussian model, Gaussian hemisphere

I. INTRODUCTION

Plane scenes are widely available in the scenes and vision computation based on planar scene is widely investigated for the past three decades. According to stratified reconstruction theories, vanishing line or the normal of a plane is crucial for affine reconstruction of a plane [1]. As a result, plane normal or its corresponding vanishing line on the image plane is important for a lot of single view based vision computation, such as metrology, localization, motion analysis, etc. It is well known that a single image doesn't provide enough information the above vision computation tasks, since the depth information of the scene is lost during the 3-D to 2-D imaging process. However, from some prior knowledge on the scene, or with sufficient geometric constraints, it is possible to perform some vision computation from a single image, e.g., computation of plane normal.

For the past three decades, a number of algorithms have been developed based on different types of scene constraints for plane normal recovery. For example, texture can be an important clue for plane normal computation. Yuchi et al and Witkan developed different algorithms to compute plane normal by using deterministic and statistical textures, respectively [2, 3]. Plane normal can also be computed from control points or lines [4, 5], designed landmarks [6, 7], planar conics [8], and other scene constraints [9, 10]. It is well known that plane normal and vanishing line can be computed from each other via the camera calibration matrix. Vanishing line or points can be computed by exploiting the parallel information in the scene [1]. Some automatic vanishing line computation algorithms can be found in [11]. Other planar constraints, like planar conics, line and conics, can also be used for vanishing line computation, which are mainly based on the circular point geometry and the absolute conic theory [1, 8~10, 12]. In practice, some geometric constraints are not always available in the scene for some specific algorithms. Moreover, some geometric constraints available in the scene can not be exploited by the type-specific algorithm. Therefore, it is important to develop a generalized computation model that can work in more general situations.

In this paper, we aim at developing a novel and generalized computation model for planar normal recovery, which can deal with different types of geometric constraints to work in more general situations. We formulate the plane normal computation problem as to maximize the likelihood that how a set of geometric constraints from the plane in the scene are satisfied. The likelihood that each constraint is satisfied is modeled by using the Gaussian model. A searching approach is then proposed to compute the plane normal by using the Gaussian hemisphere. As a result, the proposed computation model is expected to deal with different types of geometric scene constraints and is a generalized model for plane normal recovery.

II. PLANE RECTIFICATION FROM HOMOGRAPHY

Under the pin-hole camera model, a 3D point $M = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T$ in the world coordinate system is projected onto an image plane, with the image $m = \begin{bmatrix} u & v & 1 \end{bmatrix}^T$ given by [1]

$$\begin{bmatrix} u & v & 1 \end{bmatrix}^T \cong K\begin{bmatrix} R & t \end{bmatrix} \times \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \quad (1)$$

978-1-4244-7224-6 /10/$26.00 ©2010 IEEE

where the operator $\cong$ means equal up to a scale, the matrix, K represents the camera's intrinsic parameters, also called the camera calibration matrix [1]. R and t are the rotation matrix and the translation vector between the camera and the world coordinate systems.

From camera model and the projective geometry, we can derive the 2D projective geometry for a plane. Without loss of generality, Assume the reference plane is on $Z = 0$ of the world coordinate system. We can establish the relationship of a 2D point $M = \begin{bmatrix} X & Y & 1 \end{bmatrix}^T$ on $Z = 0$ and its image m from Eq. (1) as follows [1, 5]

$$m \cong \underbrace{K\begin{bmatrix} r_1 & r_2 & t \end{bmatrix}}_{H} \times \begin{bmatrix} X & Y & 1 \end{bmatrix}^T \qquad (2)$$

where r_i is the i[th] column of the rotation matrix R. Hence, M and m are related by a 3×3 full-rank matrix, called homography, in Eq. (2). The homography is a 3×3 matrix and has eight degrees of freedom. It is possible to reconstruct the homography from the vanishing line or plane normal and camera calibration matrix, according to the stratified reconstruction theory. The computation details are well investigated by in [1, 12]. Once the homography is determined, we can use it to rectify the plane, therefore to rectify the geometric attributes on the plane. Note that some image features, such as line, point, shape, etc, need to be extracted for planar geometric attribute computation.

III. THE PROPOSED ALGORITHM FOR PLANE NORMAL REOCVERY

A. *Algorithm Formulation*

We formulate the problem of plane normal computation from a set of geometric constraints as to maximize the conditional probability, which is given in the following equation

$$\arg\max_{N} P\left(N \mid C_1, C_2, \cdots C_M\right) \qquad (3)$$

where N is the normal, and $C_i\left(i = 1,2,\cdots M\right)$ is a set of geometric constraints., which are expressed as follow

$$C = \left\{C_i \mid S_i = u_i\right\}\left(i = 1,2,\cdots M\right) \qquad (4)$$

where S_i is the i[th] geometric attribute, u_i is the deterministic value, associated to the i[th] geometric attribute. Because geometric constraints can be represented in a lot of forms, such as angle, length, ratio, shape area, curvature, etc. Without loss of generality, we mainly discuss two types of fundamental geometric attributes of length ratio and angle as the geometric constraints.

Actually, it is difficult to solve Eq. (3) to compute the plane normal directly. By using Bayes' rule, we can re-arrange Eq. (3) as follows

$$P\left(N \mid C_1, C_2, \cdots C_M\right) = \frac{P\left(C_1, C_2, \cdots C_M \mid N\right)P\left(N\right)}{P\left(C_1, C_2, \cdots C_M\right)} \qquad (5)$$

where $P\left(C_1, C_1, \cdots C_M \mid N\right)$ is the likelihood, $P\left(N\right)$ and $P\left(C_1, C_2, \cdots C_M\right)$ are the prior probabilities for the normal and geometric constraints, respectively. Assume that $C_i\left(i = 1,2,\cdots M\right)$ are conditionally independent to each other, given the normal N, and $P\left(N\right)$ is constraint, e.g., a uniform distribution of N, we can derive the following equation from Eq. (5)

$$P\left(N \mid C_1, C_2, \cdots C_M\right) \propto \prod_{i=1}^{M} P\left(C_i \mid N\right) \qquad (6)$$

Hence, it shows that solving Eq. (3) is equivalent to compute the maximum likelihood in Eq. (6), given the normal direction. In other word, the solution to the normal is the one, which generates the maximum likelihood in Eq. (6).

B. *Definition of Likelihood for the Constraint*

We define $P\left(C_i \mid N\right)$ as the likelihood or probability that the i[th] constraint is satisfied, given the plane normal N. To model for the likelihood, four rules are defined:

1) The maximum probability should be obtained, where the distortion is totally removed;
2) More distortions lead to low likelihoods;
3) All geometric constraints should contribute equally to solve Eq. (6) for normal computation;
4) Geometric attributes may be in different units or different scales, which should be minimized.

To serve these purposes, we proposed using a normalized Gaussian distribution function with unit standard deviation to model the likelihood, which can meet the requirements, as specified in the four rules above. It is given as follows

$$P\left(C_i \mid N\right) = G\left(x_i\left(N\right)\right) \propto \exp\left(-\left(\frac{x_i\left(N\right) - u_i}{u_i}\right)^2 / 2\right) \qquad (7)$$

where u_i the deterministic value of the i[th] geometric constraint, as specified in Eq. (4), $x_i\left(N\right)$ is the rectified geometric attributes, given the plane normal N.

C. Searching for normal on Gaussian Hemisphere

Substitution Eq. (7) into Eq. (6) yields

$$P(N \mid C_1, C_2, \cdots C_M) \propto \prod_{i=1}^{M} G(x_i(N) \mid u_i)$$

$$\propto \prod_{i=1}^{M} \exp\left(-\left(\frac{x_i(N)}{u_i} - 1\right)^2 / 2\right) \tag{8}$$

It is well known that all the normal directions lies on the surface of Gaussian sphere, with a point representing a normal direction. For the pin-hole camera, only half of the Gaussian sphere surface, namely Gaussian hemisphere, encodes the normal directions of visible planes. In order to search the normal direction on Gaussian hemisphere, we need to first partition the Gaussian sphere into patches, with each patch representing a sampled normal. The likelihood for each sampled normal is therefore computed by using Eq. (8) based on a set of geometric constraints, as specified in Eq. (4). The maximum likelihood is thereafter searched on the Gaussian hemisphere. As a result, the corresponding normal is the final normal that we derive.

IV. Experimental Results

A. Experimental Results with Simulation Data

We first tested the performance of the proposed algorithm with simulation data. A simulated camera was used to generate all the images. The camera calibration matrix is as follows

$$K = \begin{bmatrix} 1800 & 0 & 800 \\ 0 & 1800 & 1000 \\ 0 & 0 & 1 \end{bmatrix}$$

Six length ratios and six angles were used as the basic geometric constraints to recover the plane normal direction. Among them, the six length ratios were derived from four lengths, which are 10.0, 20.0, 22.4, and 14.1 (unit in cm) on the 3D physical plane. From these four lengths, sex length ratios are obtained as 0.500, 0.447, 0.894, 0.707, 1.414, and 1.581. The six angles have the degrees of 10, 20, 30, 45, 60, and 80. A physical plane in 3D space with a normal direction of $[-0.3506 \quad 0.1580 \quad 0.9231]^T$ in the camera coordinate system is used to test the proposed algorithm. A reference frame is established from the plane with the origin at the position of $[12.0 \quad 10.0 \quad 100.0]^T$ (in cm) in the camera coordinate system.

We define N_e as the estimated normal with the proposed algorithm. N_0 is the actual normal and N_r is the optimal normal, which has the minimum angle with N_0. And we define two types of error angles to calculate the normal computation error:

$$E_{e0} = a\cos(N_e \bullet N_0),\ E_{er} = a\cos(N_e \bullet N_r)$$

In the test, random Gaussian noises with zero mean and standard deviation of 0.5 pixels were added to the image coordinates to model the practical image noises. The Gaussian hemisphere was partitioned into three resolutions of 100×200, 200×400, and 250×500. For each partition resolution, we ran one hundred trials. For each trial, we computed the plane normal and calculated E_{e0} and E_{er}, respectively. Afterwards, the means and standard deviations were calculated from the computation results of the one hundred trials. The results are shown in TABLE 1 below.

It can be observed from TABLE 1 that the standard deviations of E_{e0} and E_{er} decrease when the resolution increases, which means that high resolution can yield more reliable and robust results. Moreover, the means of E_{e0} and E_{er} also decrease with the partition resolutions. It can be observed from TABLE 1 that the computation results with the proposed model are very accurate and reliable. Actually, the partition resolution determines not only the accuracy of the computation results but also the computational complexity. For example, high resolution leads to good accuracy but high computational complexity. Hence, there should be a trade-off between them to meet the application requirements in practice.

TABLE I. Normal Computation Errors (unit in degree) under Different Partition Resolutions

Resolution	*100×200*	*200×400*	*250×500*
Mean of E_{e0}	0.4929	0.3271	0.3086
Std Dev of E_{e0}	0.2531	0.2271	0.1881
Mean of E_{er}	0.4152	0.3004	0.2883
Std Dev of E_{er}	0.4562	0.2846	0.2302
Minimum Angle	0.3475	0.1145	0.1089

We also presented the individual errors of each trial by using a partition resolution of 200×400. The results are presented in Figure 1 below. For each trial, we plotted two errors of E_{e0} and E_{er} (in degree). It can be observed that the recovered normal is equal to N_r for 30 trials, which means it gave the optimal solution (N_r) for 30 trials. It can be also observed from Figure 1 below, both E_{e0} and E_{er} for each trial are less than 1.2 degrees. Moreover, the majority (more than 90%) of E_{e0} and E_{er} is less than 0.6 degrees. The results demonstrate that the algorithm is very accurate and reliable.

B. Experimental Results with Real Image Data

The proposed algorithm was further tested using the actual image data, taken in the office, as shown in Figure 2 below. The image has a resolution of 2248×1712 (in pixel). The plane, where the white board lies (see Figure 2 below), is used as the subject for the proposed algorithm. The distance of the plane to the camera is about 6.5 meters and the height of the white board is 1.2 meters. The camera was calibrated in advance with Zhang's calibration algorithm [5]. The camera calibration matrix is as follows

$$K = \begin{bmatrix} 2356.7 & 0 & 1165.0 \\ 0 & 2356.7 & 935.8 \\ 0 & 0 & 1 \end{bmatrix}$$

We used three angles and three length ratios as the basic geometric constraints to recover the plane normal. The three angles are 90, 90, and 0 degrees, which were derived from the three visible edges of the white board (see Figure 2). The lengths are defined by the four points on the plane (see triangles in Figure 2). From these lengths, we chose three unit length ratios as the geometric constraints. From these geometric constraints, the searching approach was then applied to compute the likelihood of each normal on the Gaussian hemisphere by using a partition resolution of 200×40. As a result, a likelihood map was generated from the computation results for all normal directions (see Figure 3). The image intensity represents the probability for each normal direction, with the darker intensity for higher likelihood. The plane normal was recovered by finding the maximum likelihood, which is $\begin{bmatrix}-0.503 & 0.112 & -0.857\end{bmatrix}^T$. The corresponding position in the likelihood map is marked with a cross (see Figure 3). The actual normal was computed by using the parallel lines in the image with the algorithm proposed in [11] and the camera calibration matrix, and it is $\begin{bmatrix}-0.500 & 0.092 & -0.861\end{bmatrix}^T$. It can be calculated that the computed normal has an error of 1.2 degrees with the actual one. The results demonstrate that the algorithm is very accurate and practical.

V. CONCLUSIONS

Planar scenes are commonly found in daily life. Computation of plane normal is useful for a lot of vision computation applications, such as vision based metrology, measurement, localization, and camera calibration, etc. For the past three decades, a lot of algorithms have been developed to compute plane normal based on different types of geometric scene constraints. In this paper, we presented a novel and generalized computation model to recover plane normal by incorporating different types of geometric constraints. In this computation model, we formulated the problem of plane normal recovery from a set of geometric constraints as to maximize the likelihood that how these geometric constraints are satisfied by using the Bayes' rule. The likelihood that a geometric constraint is satisfied is modeled by Gaussian distribution function to follow the four basic rules that we defined. As a result, the maximum likelihood is computed by searching on the Gaussian hemisphere and the corresponding plane normal is the solution that we derive. The algorithm was first validated by using simulation data. The results show that the algorithm is accurate and reliable. The algorithm was further validated by using real image data, taken in an indoor office situation. It successfully computed the plane normal by using the basic constraints of length ratios and angles. The computed normal has a 1.2 degree angle with the actual normal, which demonstrates that the algorithm is accurate and practical.

ACKNOWLEDGMENT

The work presented in this paper was sponsored by a research grant from the Grant-In-Aid Scientific Research Project (No.P10049) of the Japan Society for the Promotion of Science (JSPS), Japan, and a research grant (No.L2010060) of the Department of Education, Liaoning Province, China.

REFERENCES

[1] R. Hartley and A. Zisserman, Multiple view geometry in computer vision, Cambridge, UK: Cambridge University Press, 2000

[2] A.P. Witkin, Recovering surface shape and orientation from texture, Artificial Intelligence, Vol.17, pp.17-45, 1981

[3] Yuichi Ohta, Kiyoshi Maenobu, and Toshiyuki Sakai, Obtaining surface orientation from texels under perspective projection, IJCAI, pp.746-751, 1981

[4] A. Criminisi, I. Reid, and A. Zisserman, A plane measuring device, Image Vision and Computing, Vol.17, No.8, pp. 625-634, 1998

[5] Z. Zhang, A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No. 11, pp.1330-1334, 2000

[6] Y.S. Kim, J.C.Kim, E.J.Park, and Joonwhoan Lee, Vision-based self-localization of autonomous guided vehicle using landmarks of colored pentagons, Lecture Notes in Computer Science, No. 4291, pp.133-140, 2006

[7] Kuk-Jun Yoo, Gi-Geong Jang, Sung-Ho Kim, and In-So Kweon, Fast landmark tracking and localization algorithm for the mobile robot self-localization, IFAC Workshop on Mobile Robot Technology, Jejudo Island, Korea, pp. 190-195, 2001

[8] Ip, H.H.S. , Chen, Y.S. , Planar rectification by solving the intersection of two circles under 2D homography, Pattern Recognition, Vol. 38, No. 7, pp. 1117-1120, July 2005

[9] G.H. Wang, Z.Y. Hu, F.C. Wu, Single view based measurement on space planes, Journal of Computer Science and Technology, Vol.19, No.3, pp. 374–382, 2004

[10] Feng Gao, Chellappa, R., Video metrology using a single camera, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 7, pp. 1329 – 1335, July 2010

[11] C. Rother, A new approach for vanishing point detection in architectural environments, Proc. British Machine Vision Conf., 2000

[12] D. Liebowitz, A. Zisserman, Metric rectification for perspective images of planes, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p.482, June 23-25, 1998

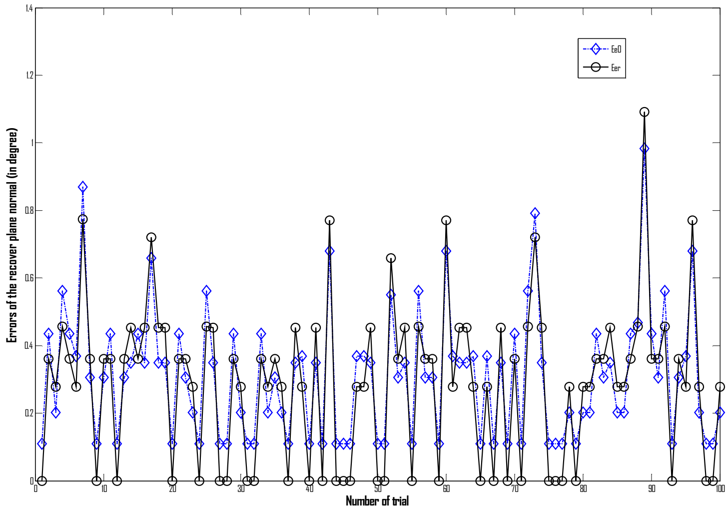

Figure 1. Plane normal computation errors with a partition resolution of *200×400*

Figure 2. Recover the normal of the plane of the white board from an indoor office image

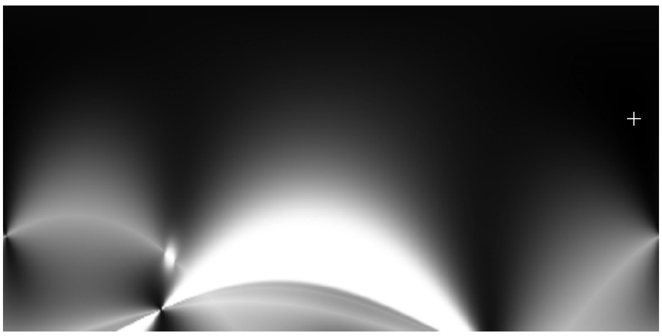

Figure 3. The likelihood map on the Gaussian hemisphere, with the computed normal marked with a cross ("+")