

A new matrix decomposition method for estimating the position of prism based stereovision

Xiaoyu Cui^{1, a}, Heting Wang¹, Yue Zhao¹, He Ma¹, Kahbin Lim²

¹ Department of SINO-DUTCH Biomedical and Information Engineering, Northeastern University, Liaoning, 110819, China

² Department of Mechanical Engineering, National University of Singapore, 119077, Singapore

^aemail: cuixy@bmie.neu.edu.cn

Keywords: Stereovision; Prism; Matrix Decomposition

Abstract. In this paper, a new matrix decomposition method was proposed for estimating the position of prism based stereovision. The prism was considered as a single optical system composed of few refractive planes. A transformation matrix which relates the coordinates of an object point to its coordinates on the image plane through the refraction of the prism was derived based on geometrical optics. A matrix decomposition method which is able to calculate the position of an arbitrary faces prism was introduced. This model further extends the application of the single-lens stereovision system using a prism to other areas. Experimentation results are presented to prove the effectiveness and robustness of our proposed model.

Introduction

Stereovision, an important field in computer vision, aims to recover the shape, size or location of an object from two or more images taken from different viewpoints. According to the number of cameras used, it can be divided into multi-camera based method and single-camera based method.

Multi-camera based method employs two or more cameras to capture images from different views. It is currently used for space exploration [1], panoramic reconstruction [2], robot navigation or other volume static target measurements and surface reconstruction [3]. Yet such measurement system has some known issues: 1) It is expensive due to the use of multiple cameras; 2) In order to obtain a high accuracy, the distance of baseline must be long enough, resulting in large system size; 3) the working condition between multiple cameras cannot be completely consistent, thus affecting the measurement accuracy; 4) cameras capture time cannot be synchronized accurately, which will cause the errors in dynamic target measurement. To overcome these limitations, the concept of single-camera stereo vision has been proposed and attracted great research interests.

The single-camera stereovision system can be classified into two categories. One is depth recovery by exploiting some known cues such as the environment or the movement of the camera [4-6]. The second is depth recovery by the reflective or refractive properties of optical devices such as mirrors [7] or prisms. Among these methods, prism-based stereovision may have the simplest system structure. With the use of only one camera, it could reduce the cost and eliminate the multi-camera synchronization problem. In addition, the system is more compact which has the potential for wide range of applications having space constraints.

In this paper, a new matrix decomposition method was proposed for estimating the position of prism based stereovision system, this new method can improve the 3D reconstruction with high precision for speed control and high dynamic quality.

Refraction of an Arbitrary Plane

The refraction of prism can be regarded as a ray refracted by two planes with inverse refractive index respectively. Thus, we will start from the virtual point formed by the refraction of an arbitrary plane, and the imaging model of prism can be acquired simply by combine the twice refractions. As shown in Fig.1, Π is an arbitrary plane with equation $n_x \cdot x + n_y \cdot y + n_z \cdot z + d = 0$ in 3D space with surface normal $N = [n_x, n_y, n_z]^T$. $A = [x_0, y_0, z_0]^T$ is a point in front of this plane. PB is a line segment which is perpendicular to plane Π . PC is an incident ray intersecting Π at point C , and the back extension line of its refracted ray intersects PB at point $P' = [x', y', z']^T$. According to Shell's law, it is clear that α is equal to the incident angle, and α' is equal to the angle of refraction. Then, the relationship between point P and P' can be expressed as below:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} qn_x^2+1 & qn_x n_y & qn_x n_z \\ qn_x n_y & qn_y^2+1 & qn_y n_z \\ qn_x n_z & qn_y n_z & qn_z^2+1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + qd \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \quad (1)$$

where $q = \tan\alpha/\tan\alpha' - 1$. From Eq. (1) we can see that the transformation from a 3D point P to its virtual image point P' is the combination of a scale transformation (with scale factor: $\tan\alpha/\tan\alpha'$) and a translation in the direction of N . When using the algebraic form, Eq. (1) can be rewritten as:

$$P' = AP + t \quad (2)$$

where $A = qNN^T + I$, $t = qdN$, and I is unit vector. Moreover, A also can be regarded as an affine transformation matrix, which is determined by q and N . Thus it has four degree of freedom (DOF) which denotes that the refraction of a plane is not a rigid transformation. According to Snell' law, q can be converted to:

$$q = \frac{n \cos \alpha'}{\cos \alpha} - 1 \quad (3)$$

where n is the refractive index. If we use unit vector V and V' to express the direction vector of CP and CP' respectively, another form of Eq. (3) is

$$q = n \frac{V' \cdot N}{V \cdot N} - 1 \quad (4)$$

Since V' also can be calculated from V , thus, when the material of the plane is determined, the q can only be influenced by the direction vector of incident ray V .

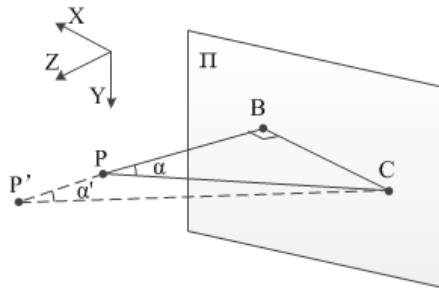


Fig.1. Ray refraction of an arbitrary plane

Perspective Projection Model

A 3D point is denoted by $P = [X, Y, Z]^T$. A 2D point is denoted by $p = [u, v]^T$. If we use \tilde{p} and \tilde{P} to denote the homogeneous coordinates: $\tilde{P} = [X, Y, Z, 1]^T$ and $\tilde{p} = [u, v, 1]^T$ the relationship between a 3D point P and its image projection p is given by

$$s\tilde{p} = M_{int}[R_e \ t_e]\tilde{P} \quad (5)$$

where s is an arbitrary scale factor, $[R_e \ t_e]$, termed the extrinsic parameters, R_e is the 3×3 rotation matrix and t_e is the 3×1 translational vector which relate the world coordinate system to the camera coordinate system, and M_{int} is the 3×3 camera intrinsic matrix. If we use superscript P^C to denote the point in camera coordinate system, then the object point in camera coordinates can be expressed as

$$\tilde{P}^C = \begin{bmatrix} R_e & t_e \\ 0^T & 1 \end{bmatrix} \tilde{P} = M_{ext} \tilde{P} \quad (6)$$

The perspective projection model of bi-prism based stereovision system is

$$\begin{cases} s\tilde{p}_l = M_{int} M_{LB} M_{ext} \tilde{P} \\ s' \tilde{p}_r = M_{int} M_{RB} M_{ext} \tilde{P} \end{cases} \quad (7)$$

Thus, the projection model of prism based stereovision is the combination of an intrinsic matrix, a scale-translate matrix and a rotate-translate matrix.

Parameters Estimation

The projection process from an object point to the image plane by prism can be formulated as the combination of an intrinsic matrix, an affine matrix and an extrinsic matrix. Since the intrinsic matrix of the camera can be calculated by single camera calibration, the calibration problem becomes to find the affine matrix and the extrinsic matrix. In Euclidean coordinates, the projection model of bi-prism based stereovision system can be written as

$$\begin{cases} s\tilde{p}_l = M_L \tilde{P} \\ s' \tilde{p}_r = M_R \tilde{P} \end{cases} \quad (8)$$

where M_L and M_R denote the perspective projection transformation of P with left and right views, which can be calculated by normal single camera calibration method, such as Direct Linear Transformation. We decompose the extrinsic matrix and acquire the following equations:

$$\begin{cases} H_L = M_{int}^{-1} M_L = M_{LB} M_{ext} \\ H_R = M_{int}^{-1} M_R = M_{RB} M_{ext} \end{cases} \quad (9)$$

If we use N_{r0} , N_{l0} and N_{b0} to express the three surface normal of the prism with no offsets, then, the surface normal of the prism can be written as

$$\begin{cases} N_{r0} = (-\sin \theta, 0, \cos \theta)^T \\ N_{l0} = (\sin \theta, 0, \cos \theta)^T \\ N_{b0} = (0, 0, 1)^T \end{cases} \quad (10)$$

In most cases, these conditions are hardly achieved in practical applications. There would be some alignment errors and the results would not be accurate if the above assumptions were used. If we use a rotation matrix R to denote the skewing of prism, then the real surface normal of each plane is

$$\begin{cases} N_l = RN_{l0} \\ N_r = RN_{r0} \\ N_b = RN_{b0} \end{cases} \quad (11)$$

According to Eq. (9), let E_R express the first three rows of H_R , then

$$E_R = A_R A_B R_e \quad (12)$$

It could also be acquired that

$$E_R E_R^T = (A_R A_B R_e)(A_R A_B R_e)^T = R A_{RBR} R^T \quad (13)$$

Since R is an orthogonal matrix, so $E_R E_R^T$ and A_{RBR} are similar matrices, then they have the same eigenvalue. Let

$$\begin{cases} E_R E_R^T = UVU^T \\ A_{RBR} = DVD^T \end{cases} \quad (14)$$

then the rotation matrix R can be solved by

$$R = UD^{-1} \quad (15)$$

Experiments and Results

The original and rectified image pairs with different offset of prism were shown in Figure 2 for visual comparison, where (a) was the original image pair, and (b) was the rectified results of the method proposed. The rotation matrix which be calculated from this three system configurations were

$$R_{sce} = \begin{bmatrix} 0.9990 & -0.0392 & 0.0226 \\ 0.0365 & 0.9932 & 0.1107 \\ -0.0268 & -0.1098 & 0.9936 \end{bmatrix}$$



Fig.2. The experimental results

As shown in these figures, the epipolar lines of original image pairs were misaligned respectively. After rectification all the re-sampled image pairs were properly rectified by the proposed method with minimal geometric distortions.

Conclusion

Prism based single-lens stereovision systems have many advantages compared with traditional two or more cameras' system. In this research, we presented a new way for calculating the position of prism, which leads to a new approach of projective rectification for stereo images. It can greatly simplify the stereo correspondence search for a bi-prism based single-lens stereovision system.

Acknowledgement

In this paper, the research was sponsored by the National Natural Science Foundation of China (Project No. 61501101), the Fundamental Research Funds for the Central Universities (N130319002), the general program of the education department of Liaoning province (L2014086) and the International cooperation projects of the technology division of Shenyang (F13-319-6-17).

References

- [1] Z.Y. Zhang, "Determining the epipolar geometry and its uncertainty: A review," *International Journal of Computer Vision*, 27(2), 161-195 (1998).
- [2] S.Kumar, C.Micheloni and C. Piciarelli, "Stereo rectification of uncalibrated and heterogeneous images," *Pattern Recognition Letters*, 31(11), 1445-1452(2010).
- [3] D.R. Wan and J. Zhou, "Self-calibration of spherical rectification for a PTZ-stereo system," *Image and Vision Computing*, 28(3), 367-375(2010).
- [4] Meijun Zhao, Kah Bin Lim, and Wei Loon Kee, Geometrical-Analysis-Based Algorithm for Stereo Matching of Single-Lens Binocular and Multi-Ocular Stereovision System, *Journal of Electronic Science and Technology*, 10(2), 107-112(2012).
- [5] Kah BinLim,WeiLoon Kee,and DaoleiWang, Virtual camera calibration and stereo

correspondence of single-lens bi-prism stereovision system using geometrical approach, *Signal Processing: Image Communication*, **28**,1059-1071(2013).

[6] Cui Xiaoyu, Lim Kahbin, Guo Qiyong and Wang Daolei, "Accurate geometrical optics model for single-lens stereovision system using a prism," *Journal of the Optical Society of America A-Optics Image Science and Vision*, **29**(9), 1828-1837 (2012).

[7] A. Fusiello, E. Trucco and A. Verri. "A compact algorithm for rectification of stereo pairs," *Machine Vision and Applications*, **12**(1), 16-22(2000).

.