3D Reconstruction using three vanishing points from a single image

Yong-In, Yoon
Chung-Ang University
221 Huksuk-Dong,
Dongjak-Ku, Seoul 156-756 South Korea
yoonyi@imagelab.cau.ac.kr
Tel. 82-2-820-5295

Jang-Hwan Im
Chung-Ang University
221 Huksuk-Dong,
Dongjak-Ku, Seoul 156-756 South Korea
jhim@cau.ac.kr
Tel. 82-2-820-5411

Dae-Hyun Kim
Chung-Ang University
221 Huksuk-Dong,
Dongjak-Ku, Seoul 156-756 South Korea
vante77@imagelab.cau.ac.kr
Tel. 82-2-820-5295

Jong-Soo Choi
Chung-Ang University
221 Huksuk-Dong,
Dongjak-Ku, Seoul 156-756 South Korea
jschoi@imagelab.cau.ac.kr
Tel. 82-2-820-5406

Abstract : This paper presents a new method which is calculated to use only three vanishing points in order to compute the dimensions of object and its pose from a single image of perspective projection taken by a camera and the problem of recovering 3D models from three vanishing points of box scene. Our approach is to compute only three vanishing points without this information such as the focal length, rotation matrix, and translation from images in the case of perspective projection. We assume that the object can be modeled as a linear function of a dimension vector V . The input of reconstruction is a set of correspondences between features in the model and features in the image. To minimize each the dimensions of the parameterized models, this reconstruction of optimization can be solved by the standard nonlinear optimization techniques with a multi-start method which generates multiple starting points for the optimizer by sampling the parameter space uniformly.

Keywords: 3D reconstruction, vanishing points, numerical optimization.

I. INTRODUCTION

Many objects such as buildings, boxes, façade and architectures involve symmetries which allow the model to be expressed with far fewer parameters[3]. We assume that the object can be modeled as a box where the coordinates of the vertices can be represented as a linear function of a dimension vector V [3,4].

Suppose the model shown in figure 1.

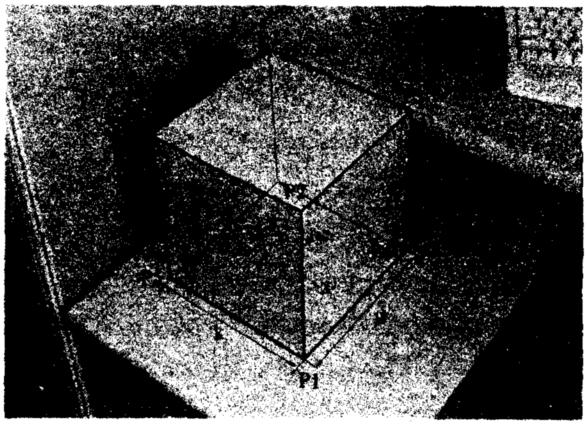

Fig. 1.An example of linearly parameterized box model

For this model, the coordinates of vertices is labeled P1, P2 ,... P7 and can be expressed as linear functions of the dimension vector $V=(k \quad w \quad h)^T$, where V is $n\times 1$ vector. Therefore, each coordinates of vertices can be represented as linear functions of parameterized vector $V=(k \quad w \quad h)^T$ in the figure 1.

$$P1=\begin{pmatrix}-k/2\\0\\h/2\end{pmatrix}=\begin{pmatrix}-0.5&0&0\\0&0&0\\0&0&0.5\end{pmatrix}\begin{pmatrix}k\\w\\h\end{pmatrix}, P2=\begin{pmatrix}-k/2\\w\\h/2\end{pmatrix}=\begin{pmatrix}-0.5&0&0\\0&1&0\\0&0&0.5\end{pmatrix}\begin{pmatrix}k\\w\\h\end{pmatrix}..$$

In this case the parameter vector V will refer to dimensions that are only meaningful when positive. In figure 1, if the positions of vertices is eight numbers, only three parameters can be characterized. For this model, it is possible to recover the dimensions from measurements in a single image.

The input of reconstruction is a set of correspondences between features in the model, lines and points, and features in the image. Therefore, we determine a proper projection model for the dimensions of the object and its pose acquired with a camera in the case of perspective projection. To minimize each the dimensions of the parameterized models, this reconstruction of optimization can be solved by the standard nonlinear optimization techniques with a multi-start method which generates multiple starting points for the optimizer by sampling the parameter space uniformly.

David and Camillo[4] described effective method for recovering the dimensions of object and its pose from a single image acquired under orthographic and perspective projection models. However, in the case of perspective projection the dimensions of vector can be determined to compute only the focal length of camera, rotation matrix and translation using the vanishing points. In this paper, we propose a new method which is calculated to use only three vanishing points in order to compute the dimensions of object and its pose from a single image taken by a camera and the problem of recovering 3D models from three vanishing points of box scene without this information such as the focal length, rotation matrix, and translation from images in the case of perspective projection.

This paper can be represented as follows: Section 2 presents the solution of the proposed reconstruction algorithm in the case of perspective projection, finding vanishing points and the texture mapped 3D model. Section 3 describes result obtained with the proposed algorithm on actual images and on simulated and the result of the texture mapped 3D model. Section 4 presents conclusions and future work of this paper.

2. Reconstruction Algorithm

2.1 Finding vanishing points

An input image is constrained information that can be obtained with image because of considering only an image taken by camera in the case of perspective projection. Therefore, we can extract the information of a vanishing point to have orthogonal property from images. Many systems using vanishing points from image have been proposed[1]. In this paper, the constraints which can be used are the model of an image acquired under perspective projection model and exploiting three vanishing points.

In this paper, we also exploit to compute the dimension of model using the vanishing points[6]. A vanishing point corresponds to the projection of the intersection of parallel lines at infinity in the X, Y, and Z directions from the model of an image. To explore very accurately the vanishing point, the user defines interactively lines of the model and finds correspondences between edges in the model and edges in the image by selecting a line in the model. To compute the vanishing points of X, Y, and Z directions, the user selects n lines in the model that are parallel to the X, Y, and Z axis of the object. Let $L_1, L_2, \ldots, L_n$ be a set of image lines corresponding to parallel to the X axis and let V_x be the vanishing points defined by these lines of the X direction. Then, the homogeneous coordinates of the vanishing points in the X, Y, and Z directions are the vectors V_x, V_y, V_z that minimize $\sum \left\| L_i^T V_x \right\|^2, \sum \left\| L_i^T V_y \right\|^2, \sum \left\| L_i^T V_z \right\|^2$. These vectors can be computed by eigenvalue decomposition of $D^T D$, where D is the matrix whose rows consist of the L_i^T's, using a constrained optimization from linear algebra[2]. This techniques can be computed to the vanishing points of the X, Y, and Z directions. The result of computing the vanishing points from the model of an image is represented figure 2.

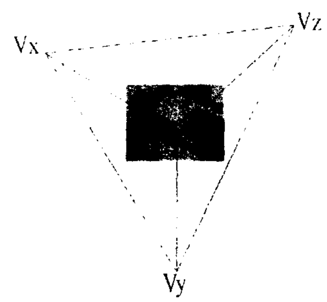

Figure 2. Each vanishing points for the model of X, Y, and Z directions

2.2 The Proposed Recovering Scene Dimensions

In the case of perspective projection model the homogeneous camera projection matrix P, which relates coordinates of points in the model to their projections on the image plane, can be given by:

$$K=\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}, R=\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{22} & r_{33} \end{bmatrix}, T=\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}, C=\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$P=K[R \quad T]$$

where K is the intrinsic parameter, f denotes the focal length of the camera and the principal points and R and T are the rotation and translation of the camera with respect to the model frame.

In order to recover the dimensions vector V, it is used to only three vanishing points. According to the box model, the coordinates of the jth vertex in the world frame are given by P_j.

Let $L_{jk}=\left(L^x_{jk} \quad L^y_{jk} \quad L^z_{jk}\right)^T$ show the homogenous coordinates of the line in the image plane connecting points j and k. Then, the constraint that the projection of the jth vertex in the image should lie along this line can be represented as follows:

$$L^T_{jk}P_jC=0 \qquad (1)$$

In the equation, consider we are given V_x, V_y and V_z. Then, we get the following proportions:

$$V_x = KR\hat{X}, V_y = KR\hat{Y}, V_z = KR\hat{Z}$$

where $\hat{X}, \hat{Y}$ and $\hat{Z}$ are the unit vectors along the X, Y, and Z axes respectively.

If V_x, V_y and V_z are substituted for the equation (1), then we can be expressed as follows:

$$L^T_{jk}K(R+T)C = L^T_{jk}(KR_1X + KR_2Y + KR_3Z + KT)$$
$$= L^T_{jk}(V_xX + V_yY + V_zZ + KT) = 0$$
$$\Rightarrow L^T_{jk}\begin{bmatrix} V_x & V_y & V_z & K \end{bmatrix}\begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix} = 0$$

To minimize the vector dimensions of the parameterized models, setting K=1, we compute the vector dimensions. We know that we have no effect on this reconstruction of optimization.

Finally the above equation can be obtained because of computing the dimensions vector of the model. Let A be a matrix formed by stacking the rows of the form

$$L^T_{jk}\begin{bmatrix} V_x & V_y & V_z & K \end{bmatrix}.$$

This equation can be obtained by the constrained optimization method[2]

2.3 3D reconstruction And Texture mapping

In the previous section, we can determine the camera projection matrix P. From this information, the 3D point positions can be found by the image of back-projection. Therefore, these points are used an image point triangulation to obtain 3D structure. This structure is rendered using a standard texture mapping procedure. Then we determine the model of 3D structure.

3. EXPERIMENTAL RESULTS

In this experiment, the perspective reconstruction method was used. This technique does make use of vanishing point information and formulated as optimization.

The result was obtained using photographs taken with Canon EOS D30 digital camera. The image size was 2160×1440 in the high-resolution mode.

In order to prove the proposed reconstruction system, the experimental image is used to simple image of a box. Figure 3a shows a box model of image and Figure 3b represent texture-mapped reconstruction of the model. The vector V, which gives the dimensions of the box, were measured by hand and found to be (in cm)(20 20 20)^T(Fig.3a),(30 20 10)(Fig.4), and (30 20 7)(Fig.5). After choosing an appropriate scaling factor, the previous method[4] gave an estimate (in cm) of (18.64 21.8 18.8)^T (Fig.3a),(31.2 14.9 9.5)^T(Fig.4), and (30.6 15.46 8.1)^T(Fig.5) and the proposed reconstruction an estimate in centimeters of (21.2 20.46 18.1)^T (Fig.3a),(28.3 19.25 14.96)^T (Fig.4), and (28.64 19.78 7.7)^T(Fig.5).

4. CONCLUSIONS AND FUTURE WORKS

We have presented a new method which is calculated to use only three vanishing points in order to recover the dimensions of object and its pose from a single image of perspective projection. Experimental results show that our method have the accuracy and efficacy of the dimensions vector in comparison with the actual images. Future work will describe how to use objects to recover the use of automated edge extraction.

Acknowledgment

This research was supported by the Ministry of Education, Seoul, Korea, under the BK21 project, and by the Ministry of Science and Technology, Seoul, Korea, under the NRL project.

REFERENCE

[1] B. Caprile and V. Torre, Using vanishing points for camera calibration, *IJCV*, pp. 127-140, 1990.

[2] O. Faugeras, Three-dimensional computer Vision, MIT Press, 1993.

[3] R P.E Debevec, C.J Taylor, and J. Malik, Modeling and Rendering Architecture from Photographs: A Hybrid Geometry and Image-Based Approach, In *Proceedings of ACM SIGGRAPH 1996*, ACM Press / ACM SIGGRAPH, pp.11-21,Aug. 1996

[4] D. Jelinek and C.J. Taylor, Reconstruction of linearly parameterized models from single images with a camera of unknown focal length, IEEE PAMI vol.23, no.7 July 2001

[5] Todd K. Moon and Wynn C. Stirling, Mathematical Methods and Algorithms for signal processing, Prentice Hall, 2000

[6] R.T Collins and R.S Weiss, Vanishing point calculation as a statistical inference on the unit sphere, In Proc.Third Int. Conference on Computer Vision, pp. 400-403,Dec.1990

[7]O.D Faugeras. Stratification of three-dimensional vision: projective, affine, and metric representation. J. Opt. Soc. Am., A12: 465-84,1995

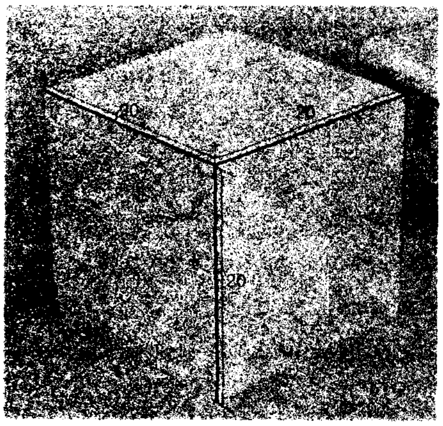

Figure 3a : A box(square) with weak perspective.

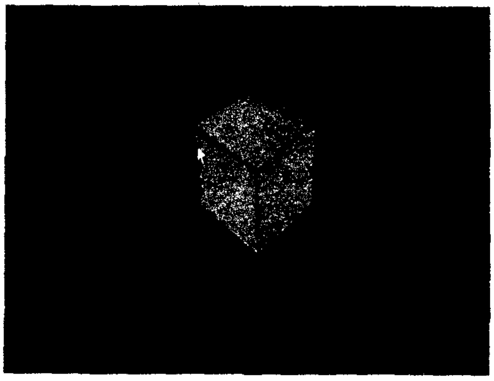

Figure 3b: Texture mapped reconstruction of the scene model 1

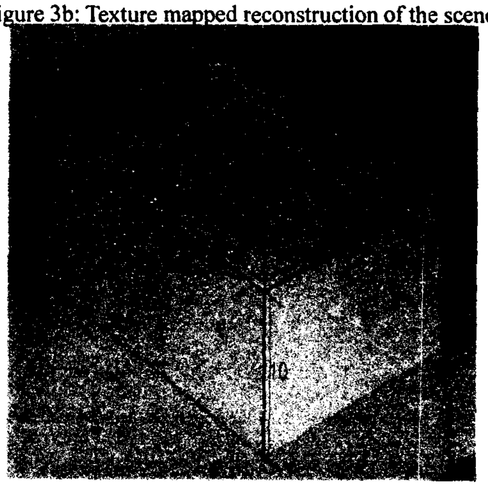

Figure 4: A box(rectangular) with weak perspective

Figure 5: A box(rectangular) with weak perspective