MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

Self-Calibration for Metric 3D Reconstruction Using Homography

Cheng-Yuan Tang[1], Yi-Leh Wu[2], Pei Ching Hu[1], Hsien-Chang Lin[3], Wen-Chao Chen[3]

[1] Department of Information Management, Huafan University, Taipei, Taiwan
[2] Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan
[3] Electronics and Optoelectronics Research Laboratories, ITRI, Hsinchu, Taiwan
e-mail: cytang@cc.hfu.edu.tw

Abstract.

In this paper, our goal is to reconstruct metric 3D models for large structures (or other 3D objects) from uncalibrated images. The internal camera parameters are estimated by using homographic constraints. More specifically, we propose to use straight lines and vanishing points to compute the infinite homography. However, because vanishing points are very sensitive to noise, we propose further constraints on homography with multiple image views. After obtaining the internal camera parameters, the metric 3D models can then be accurately reconstructed. Some preliminary results are presented in this paper.

1. Introduction

Before reconstruct 3D models from images captured by cameras, camera calibration plays an essential role. Camera calibration has long been an important research topic in computer vision. Camera calibration is the process that models the mathematic relationship between the 3D coordinates of objects in a scene and their 2D coordinates on the projected images. The parameters of a camera are classified into two categories: internal camera parameters and external camera parameters. Internal camera parameters define the properties of the geometrical optics and the external camera parameters define the rotation and translation of the camera.

There are two main methods to obtain metric 3D reconstruction from uncalibrated images: (a) using the transformation matrix or (b) using self-calibration techniques. The transformation matrix can be obtained by manually measuring the target objects. However, it is usually difficult to measure in metric for large structures. Self-calibration is the process that determines internal camera parameters directly from uncalibrated images. Once the process has been completed, it is possible to obtain metric 3D reconstruction directly from images. There exist many literatures on self-calibration; e.g., Pollefeys and Gool [2], Faugeras [4], Zhang and Faugeras [5], Hartley and Zisserman [1], Huang and Chen [9]. During the self-calibration process, two robust estimators, the least median of squares (LMedS) [7][8], and the random sample consensus (RANSAC) [3][7], are commonly used to resist the outliers.

In this paper, we propose techniques that employ the homography constraint for camera self-calibration and internal camera parameters estimation. Some preliminary results in metric 3D reconstruction are shown in the experiment section.

2. Camera Geometry and Camera Model

We use the projective geometry throughout this paper to describe the perspective projection of the 3D scene onto 2D images. This projection is described as follows:

$$\mathbf{x} = \mathbf{PX} \tag{1}$$

where $\mathbf{P}$ is a 3×4 projection matrix that describes the perspective projection process, $\mathbf{X} = [X, Y, Z, 1]^T$ and $\mathbf{x} = [x, y, 1]^T$ are vectors containing the homogeneous coordinates of the 3D world coordinate, respectively, 2D image coordinate.

When the ambiguity on the geometry is metric, (i.e., Euclidean up to an unknown scale factor), the camera projection matrices can be put in the following form:

$$\mathbf{P} = \mathbf{K}[\mathbf{R} \mid -\mathbf{Rt}] \tag{2}$$

with $\mathbf{t}$ and $\mathbf{R}$ indicating the position and orientation of the camera and $\mathbf{K}$, an upper diagonal 3×3 matrix containing the intrinsic camera parameters:

$$\mathbf{K} = \begin{bmatrix} f_x & s & u_x \\ & f_y & u_y \\ & & 1 \end{bmatrix} \tag{3}$$

where f_x and f_y represent the focal length divided by the horizontal and vertical pixel dimensions, s is a measure of the skew, and (u_x, u_y) is the principal point.

2.1. Two-View Geometry

Consider two image points $\mathbf{x}$ and $\mathbf{x}'$ are projected from a 3D point $\mathbf{X}$ observed by two cameras with optical centers $\mathbf{O}$ and $\mathbf{O}'$, these five points form a common plane called the epipolar plane. The points $\mathbf{e}$ and $\mathbf{e}'$ are called the epipoles of the two cameras where the epipole $\mathbf{e}'$ is the projection of the optical center $\mathbf{O}$

of the first camera in the image observed by the second camera and vise versa. If $\mathbf{x}$ and $\mathbf{x}'$ are projection of the same point, then $\mathbf{x}'$ must lie on the epipolar line associated with $\mathbf{x}$, hence the epipolar constraint. The epipolar constraint plays an important role in stereo vision analysis. When the intrinsic parameters of the cameras are known, the epipolar constraint can be represented algebraically by a 3x3 matrix, called the essential matrix. Otherwise, the epipolar constraint represented by a 3x3 matrix is called the fundamental matrix, **F**.

2.2. Homography

There exists a relationship between the points from two images shooting from different viewing angles if the points lie on the same 3D plane (as shown in Figure 1). This relationship can be represented as a 3×3 transformation matrix, the planar homography matrix **H,** as follows:

$$\mathbf{x}' = \lambda \mathbf{H}\mathbf{x} \tag{4}$$

where λ is a scalar

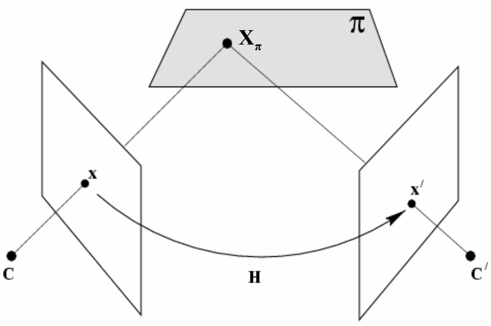


Figure 1. The homography induced by a plane.

Expanding above equation, we have

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{5}$$

If we have n matched point pairs from the same 3D plane, the above equation can be used to solve the 3x3 homography matrix **H** by applying the SVD method [11].

2.3. Infinite Homography

The absolute conic is an imaginary conic situated on the plane at infinity. It corresponds to the equation $X^2 + Y^2 + Z^2 = 0$ and $T = 0$ (for points $\mathbf{X} = [X \quad Y \quad Z \quad T]^T$).

If the camera does not change, then the image of the absolute conic $\boldsymbol{\omega}$ and its dual $\boldsymbol{\omega}^*$ will also stay the same for all views, as shown in Figure 2. These are represented by the following matrices:

$$\boldsymbol{\omega} = \mathbf{K}^{-T}\mathbf{K}^{-1} \tag{6}$$

$$\boldsymbol{\omega}^* = \mathbf{K}\mathbf{K}^T \tag{7}$$

with $\mathbf{K}$ the upper triangular matrix containing the internal camera parameters.

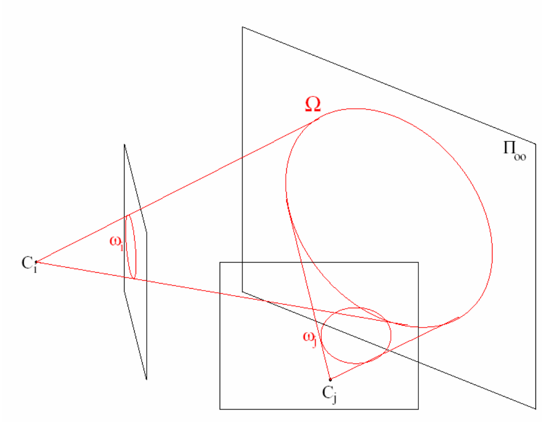


Figure 2. An absolute conic on the infinity plane.

The constraint that the dual image of the absolute conic should be the same for all the views can then be expressed as follows:

$$\lambda \mathbf{K}\mathbf{K}^T \propto \mathbf{H}_{\infty ij}\mathbf{K}\mathbf{K}^T\mathbf{H}^T\mathbf{H}_{\infty ij} \tag{8}$$

Scaling $\det \mathbf{H}_{\infty ij}$ to one eliminates the scale factor λ. Therefore, once $\mathbf{H}_{\infty ij}$ is known, Equation 8 represents a set of linear equations from which the elements of $\mathbf{K}\mathbf{K}^T$ can be obtained. The internal camera parameters can then be obtained through Cholesky factorization.

From Equation 8, we should compute the infinite homography before we estimate the internal camera parameters. There are two methods to compute the infinite homography. The first method requires three corresponding vanishing points and the fundamental matrix, and the second method requires one corresponding vanishing point, one vanishing line and the fundamental matrix [1].

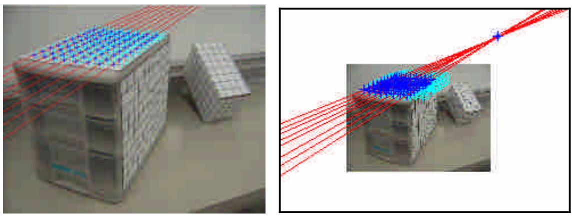
3. Self-Calibration

Self-calibration is the process that computes the metric properties of the camera and/or the scene from a set of uncalibrated images. The key concept for self-calibration is the absolute conic discussed in Section 2.3. The absolute conic stays constant under all rigid transformations of space and it encodes the metric structure of the scene (i.e., Euclidean structure up to scale λ).

3.1. Computing Vanishing Points and Lines

In order to compute the infinite homography, we need to accurately compute the vanishing points and vanishing lines. Unfortunately, the process to compute the vanishing points and the vanishing lines is very sensitive to noise (outliers). Therefore we employ a robust estimator, LMedS [7][8], to compute the vanishing points. The procedure to compute the vanishing points is described as follows:

1. Select multiple points that form a line.
2. Compute the line using LMedS. As shown in Figure 3(a), 10 points are used to fit one line.
3. Compute the intersection of lines from Step 2 with LMedS. As shown in Figure 3(b), the vanishing point is determined by 8 lines.



(a) (b)

Figure 3. A vanishing point.

The vanishing line can be determined by two or more vanishing points using LMedS. After the vanishing points or vanishing lines are obtained, the infinite homography can be estimated by Equation 8.

3.2. Computing internal camera parameters

For simplicity, let the internal camera parameters be constant, we set the aspect ratio to 1, the skew to 0, and the principle point to $(0,0,1)^T$ [9],. Once the infinite homography is estimated, the stratified algorithm [2], $\mathbf{KK}^T \propto \mathbf{H}_{\infty 1i}\mathbf{KK}^T\mathbf{H}_{\infty 1i}^T$, can be used to compute $\mathbf{K}$.

$$\omega^* = \lambda KK^T = \lambda \begin{bmatrix} f^2 & & \\ & f^2 & \\ & & 1 \end{bmatrix}$$

$$KK^T = \lambda H_{\infty 1i} KK^T H_{\infty 1i}{}^T$$

$$\begin{bmatrix} f^2 & & \\ & f^2 & \\ & & 1 \end{bmatrix} = \lambda \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} f^2 & & \\ & f^2 & \\ & & 1 \end{bmatrix} \begin{bmatrix} h_{11} & h_{21} & h_{31} \\ h_{12} & h_{22} & h_{32} \\ h_{13} & h_{23} & h_{33} \end{bmatrix} \quad (9)$$

Here, i ($i = 2,...,n$) denotes the index of images. For example, $\mathbf{H}_{\infty 12}$ represents the infinite homography from the 1st image and the 2nd image. Equation 9 is then used to estimate the focal length f.

3.3. Constraints with Multiple Views

The internal camera parameters will not be accurate when the infinite homography $\mathbf{H}_{\infty 1i}$ is not accurate enough. We propose to apply some constraints to filter out singular cases while estimating the infinite homography $\mathbf{H}_{\infty 1i}$.

The relationship of the homography among multiple image views can be depicted as in Figure 4.

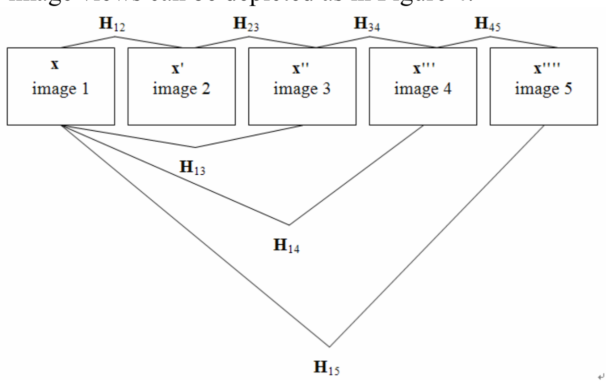


Figure 4. The relationships of homography among five images.

In Fugure 4, the corresponding 2D coordinates of a fixed 3D point $\mathbf{X}$ projected in five images are $\mathbf{x}, \mathbf{x}', \mathbf{x}'', \mathbf{x}'''$ and $\mathbf{x}''''$. The homography from the 1st image to the 2nd image is denoted as $\mathbf{H}_{12}$; i.e., $\mathbf{x}' = \mathbf{H}_{12}\mathbf{x}$. Similarly, the homography $\mathbf{H}_{23}, \mathbf{H}_{34}, \mathbf{H}_{45}, \mathbf{H}_{13}, \mathbf{H}_{14}, \mathbf{H}_{15}$ can be defined as follows:

$$\mathbf{x}' = \mathbf{H}_{12}\mathbf{x}\text{, } \mathbf{x}'' = \mathbf{H}_{23}\mathbf{x}'\text{, } \mathbf{x}''' = \mathbf{H}_{34}\mathbf{x}''\text{, } \mathbf{x}'''' = \mathbf{H}_{45}\mathbf{x}'''$$

or

$$\mathbf{x}'' = \mathbf{H}_{23}\mathbf{H}_{12}\mathbf{x}$$

$$\mathbf{x}''' = \mathbf{H}_{34}\mathbf{H}_{23}\mathbf{H}_{12}\mathbf{x}$$

$$\mathbf{x}'''' = \mathbf{H}_{45}\mathbf{H}_{34}\mathbf{H}_{23}\mathbf{H}_{12}\mathbf{x}$$

The result is that we can obtain additional constraints from the formation of the homography matrices from multiple views; e.g.,

$$\mathbf{H}_{13} = \mathbf{H}_{23}\mathbf{H}_{12}$$

$$\mathbf{H}_{14} = \mathbf{H}_{34}\mathbf{H}_{23}\mathbf{H}_{12}$$

$$\mathbf{H}_{15} = \mathbf{H}_{45}\mathbf{H}_{34}\mathbf{H}_{23}\mathbf{H}_{12}$$

These resulting constraint equations can help to filter out some singular cases while estimating the infinite homography.

4. Experimental Results

The first experiment employs the stratified algorithm to estimate the internal camera parameters. First, the vanishing points are obtained. Then the infinite homography is obtained by using the resulting vanishing points. Finally, the internal camera parameters are computed from the infinite homography. In reality, the points on the vanishing lines in images can be detected automatically. For simplicity, points on the vanishing lines in images are detected manually in this experiment. Figure 5 shows the original images and the results after detecting vanishing points and vanishing lines (in blue). The infinite homography can then be computer by Equation 8.

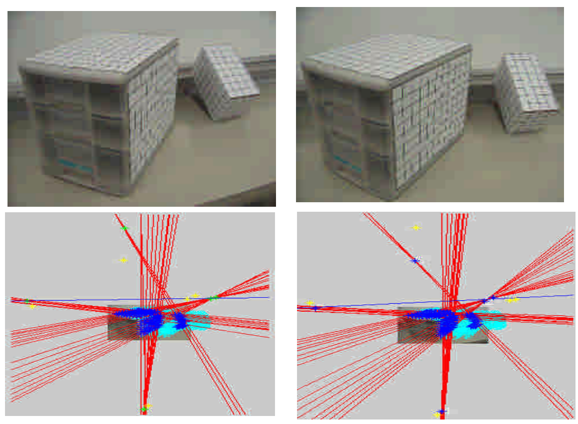


Figure 5. Computing the intersections of lines, vanishing points and vanishing lines.

As discussed in Section 3.2, for simplicity, we set the aspect ratio to 1, skew to 0, and the principle point to $(0,0,1)^T$. After computing the infinite homography, the internal camera parameter is obtained by Equation 9 as the estimated focal length to be 892.3304.

Figure 6. Two views in a controlled environment.

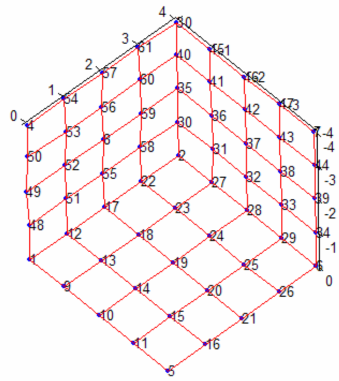


Figure 7. Reconstruction result.

Figure 6 shows two images of the same cube with markings from different angles in a controlled environment. The corresponding points are depicted in blue. The corresponding metric 3D reconstruction is shown in Figure. 7.

Figure 8. Two views of a structure.

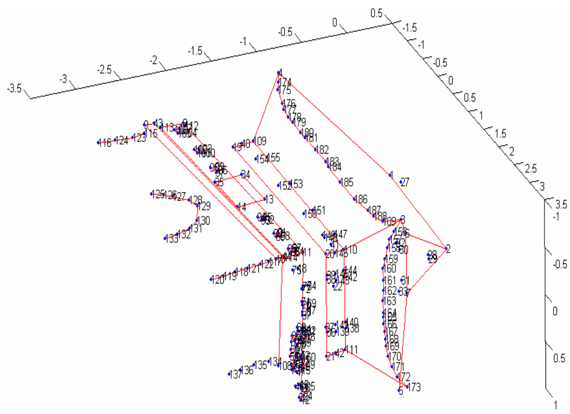


Figure 9. Metric 3D reconstruction from 2 views.

Figure 8 shows two images of the same structure from different angles. The corresponding points are depicted in blue. The corresponding metric 3D reconstruction is shown in Figure 9.

5. Conclusion

The goal of this paper is to reconstruct metric 3D models for large structures from uncalibrated images. We propose to use straight lines and vanishing points to compute the infinite homography and then estimate the internal camera parameters. However, because vanishing points are very sensitive to noise, we propose further constraints on homography with multiple image views. After obtaining the internal camera parameters, the metric 3D models can then be accurately reconstructed. Our prelimenary experimental results show that although we try to employ some existing robust methods to overcome the noise/outliers problem, the estimation of the internal camera parameters is still very sensitive to noise. Our future woks focus on improving the techniques for more robust and stable estimation in camera self-calibration.

Acknowledgement

This work was partially supported by the National Science Council, Taiwan, under the Grants No. NSC95-3114-P-001-002-Y02, NSC95-3114-P-001-001-Y02 (the iCAST project) and NCS95-2221-E-011-059.

References

[1] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[2] M. Pollefeys and L. Van Gool, "Stratified Self-Calibration with the Modulus Constraint," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, No. 8, pp. 707-724, 1999.

[3] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2003.

[4] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, 1993.

[5] Z. Zhang and O. Faugeras, *3D Dynamic Scene Analysis: A Stereo Based Approach*. Springer-Verlag, 1992.

[6] A. Fitzgibbon and A. Zisserman, "Automatic 3D Model Acquisition and Generation of New Images from Video Sequences," *Proceedings of European Signal Processing Conference (EUSIPCO '98)*, Greece, pp. 1261-1269, 1998.

[7] P. Meer, D. Mintz, D. Kim, and A. Rosenfeld, "Robust Regression Methods for Computer Vision: A Review," *International Journal of Computer Vision*, Vol. 6, pp. 59-70, 1991.

[8] P. J. Rousseeuw, "Least Median of Squares Regression," *Journal of the American Statistical Association*, Vol. 79, pp. 871-880, 1984.

[9] C.-R. Huang, C.-S. Chen, and P.-C. Chung, "An Improved Algorithm for Two-Image Camera Self-Calibration and Euclidean Structure Recovery Using Absolute Quadric," *Pattern Recognition,* pp. 1713-1722, 2004.

[10] M. Bénallal and J. Meunier, "Camera Calibration with Simple Geometry," *Moroccan Journal of Control Computer Science and Signal Processing*.

[11] C.-Y. Tang, H.-L. Chou, Y.-L. Wu and Y.-H. Ding, "Fast and Robust Algorithm Using Coplanar Constraints to Estimate Fundamental Matrices," *2006 IEEE International Conference on Systems, Man and Cybernetics (SMC2006).*