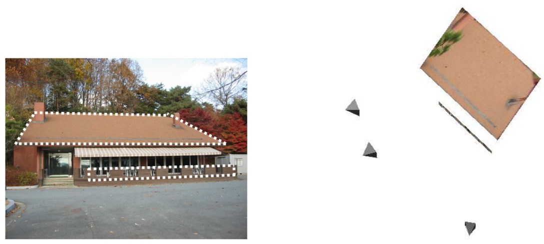
Camera Calibration and Scene Modeling from Arbitrary Parallelograms Imposing the Multiview Constraints

Jae-Hean Kim
gokjh@etri.re.kr

Bon-Ki Koo
bkkoo@etri.re.kr

Content Research Division
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea



(a) (b)

Figure 1: (a) Two parallelograms used as the input for the proposed algorithm. (b) Reconstructed model and camera pose

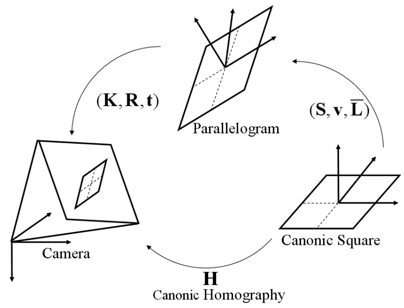


Figure 2: The projection of the canonic square onto the vertices of the imaged parallelograms.

This paper proposes a novel framework using geometric information on parallelograms for camera calibration and scene modeling (Fig 1). In the images captured in man-made environments, there are many primitives giving the geometric information. Among primitives, parallelograms and parallepipeds are frequently present in the scene, such as the architecture. The scene's affine structure is embedded in them and their Euclidean structure related to the shape can be used to upgrade the affine structure to the metric one.

Wilczkowiak *et al*. suggested an elegant formalism using parallelepipeds of which at least the six vertices are visible in views [3]. This method is a factorization-based approach that computes the camera parameters and scene structure parameters linearly and simultaneously in one step. Due to the multiview constraints imposed on a factorization approach, it is possible to obtain the consistency of rigid transformations among cameras. If this consistency is not guaranteed, the estimation accuracy can be degraded. It has been known that all image measurements should be used simultaneously to obtain optimal estimates as in the factorization approach [2]. The primitive that can give full affine information is not only a parallelepiped. Parallelograms are more general primitives in man-made environments. Since pairs of parallelograms do not always form two faces of a parallelepiped, the methods using parallelograms are more flexible in use [1]. However, the previous linear approaches using parallelograms cannot guarantee the above consistency because they are based on the information extracted from the individual pairs of camera images and should combine the individual results when more than two views are given.

In this paper, we suggest a factorization-based framework that utilizes parallelograms in general position and ensures consistent results. The general position means that the parallelograms need not to be the part of a parallelepiped. One measurement matrix includes all image measurement in all views and is factorized into the camera and plane parameters. The contributions of this paper concerns the formulation of the measurement matrix from parallelograms and the computation of the scale factors necessary for the factorization.

Let $\mathbf{x_s}$ be homogeneous coordinates of vertex of the canonic square. Consider the projection of a parallelogram's vertices into a camera image plane. The projection of the corresponding vertex in the image is:

$$\begin{aligned} \mathbf{x} &\cong \mathbf{K}\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}\begin{bmatrix} \bar{\mathbf{S}} & \mathbf{v} \\ \mathbf{0}^T & 1 \end{bmatrix}\begin{bmatrix} \bar{\mathbf{L}} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}\mathbf{x_s} \\ &= \mathbf{MNx_s} = \mathbf{Hx_s}. \end{aligned}$$

The 3×3 matrix $\mathbf{K}$ is the camera calibration matrix. The 3×4 matrix $\mathbf{M}$ encapsulates the camera intrinsic and extrinsic parameters. $\bar{\mathbf{S}}$ is the 3×2 submatrix consisting of the first two columns of a rotation matrix $\mathbf{S}$ representing the parallelogram's orientation and a vector $\mathbf{v}$ is its position. The matrix $\bar{\mathbf{L}}$ represents the parallelogram's shape (intrinsic paramter).

The matrix $\mathbf{H}$ will be called the *canonic homography*. It represents a perspective projection that maps the vertices of the canonic square onto the vertices of the imaged parallelograms (Fig 2). Given image points for four vertices, the canonic homography can be computed up to scale, even though we do not know prior knowledge on intrinisic or extrinisic parameters.

We may gather the canonic homography $\tilde{\mathbf{H}}$ estimated up to scale for all m cameras and n parallelograms into the following single matrix:

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{H}}_1^1 & \ldots & \tilde{\mathbf{H}}_1^n \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{H}}_m^1 & \ldots & \tilde{\mathbf{H}}_m^n \end{bmatrix}. \tag{1}$$

The matrix $\tilde{\mathbf{W}}$ will be called the *measurement matrix*. When the scale factors λ_i^j are recovered, the measurement matrix can be factorized:

$$\begin{aligned} \begin{bmatrix} \lambda_1^1\tilde{\mathbf{H}}_1^1 & \ldots & \lambda_1^n\tilde{\mathbf{H}}_1^n \\ \vdots & \ddots & \vdots \\ \lambda_m^1\tilde{\mathbf{H}}_m^1 & \ldots & \lambda_m^n\tilde{\mathbf{H}}_m^n \end{bmatrix} &= \begin{bmatrix} \mathbf{H}_1^1 & \ldots & \mathbf{H}_1^n \\ \vdots & \ddots & \vdots \\ \mathbf{H}_m^1 & \ldots & \mathbf{H}_m^n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_m \end{bmatrix}\begin{bmatrix} \mathbf{N}^1 & \ldots & \mathbf{N}^n \end{bmatrix}. \end{aligned} \tag{2}$$

However, since the canonic homographies are obtained up to scale, the measurement matrix cannot be factorized as its current form. In this paper, it will be described how to obtain the scale factors using the *relative homography* known to be a planar homology having the form $\mathbf{I}+\mathbf{ab}^T$, where $\mathbf{a}$ and $\mathbf{b}$ are arbitrary 3-vectors [4]. The results obtained from the factorization are equivalent to 3D reconstruction up to affine transformation. We can upgrade the results to metric ones using metric constraints from scene geometry and camera intrinsic parameters [1]. For example, the orthogonality of a parallelogram's edges and zero skew of the intrinsic parameters can be used.

[1] Jae-Hean Kim. Linear stratified approach for 3D modelling and calibration using full geometric constraints. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 2144 – 2151, Miami, FL, USA, June 2009.

[2] C. Tomasi and T. Kanade. Shape and motion from image streams under orthogrphy: a factorization method. *International Journal of Computer Vision*, 9(2):137–154, November 1992.

[3] Marta Wilczkowiak, Peter Sturm, and Edmond Boyer. Using geometric constraints through parallelepipeds for calibration and 3D modelling. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(2):194–207, February 2005.

[4] Lihi Zelnik-Manor and Michal Irani. Multiview constraints on homographies. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(2):214–223, February 2002.