Linear Stratified Approach for 3D Modelling and Calibration using Full Geometric Constraints

Jae-Hean Kim
Electronics and Telecommunications Research Institute (ETRI)
161, Gajeong-dong, Yuseong-gu, Daejeon, 305-350, Republic of Korea
gokjh@etri.re.kr

Abstract

There have been many approaches to obtain 3D modeling and camera calibration simultaneously from uncalibrated images using parallelism, orthogonality and self-calibration constraints. These approaches can give more stable results with fewer images and allow us to gain the results with only linear operations in most cases. It has been proved that the estimation results are accurate enough to be used as the initial values for nonlinear optimization to refine the results. In this paper, it is shown that all the linear constraints used in the previous works performed independently up to now can be implemented easily in the proposed linear method. The proposed method uses a stratified approach, in which affine reconstruction is performed first and then metric reconstruction. In this procedure, the additional constraints newly extracted in this paper have an important role for affine reconstruction in practical situations. The study on the situations that can not be dealt with by the previous approaches is presented and it is shown that the proposed method being able to handle the cases is more flexible in use.

1. Introduction

To develop a method calibrating cameras and reconstructing 3D model simultaneously from images are one of the important topics in the field of computer vision. There have been many approaches concerning camera calibration and 3D model reconstruction.

First, the classical approach uses a calibration rig of which the dimension is given a priori [14], [17]. However, this approach is inconvenient in that a heavy and large calibration rig is required. Second, the camera can be calibrated from solely the image sequences using only the constraints on the internal and/or external parameters. This approach is known as auto-calibration and provides great flexibility [12], [11]. However, to acquire high quality, a lengthy image sequence and a relatively accurate feature correspondence set are necessary. Third, the information from restricted camera motion can be used to calibrate camera or 3D reconstruction [10], [4], [1]. Finally, there have been many methods that use the properties of the scene or special objects [2], [8], [6], [13], [15], [3], [16], [7]. As these methods do not require a calibration rig but use scene geometry or the shapes of objects as the calibration rig, accurate calibration results can be acquired without lengthy image sequences. The proposed algorithm is categorized into these methods.

In this paper, we propose linear stratified approach that can use full geometric constraints. The method can obtain affine structure from parallelograms in general position and all available geometric information including self-calibration constraints can be utilized to upgrade the result to a metric one. The parallelograms need not to be the part of a parallelepiped. To cope with the general position of parallelograms in practical situations, we extract additional constraints which are from parallelograms and have an important roll for an affine reconstruction. It can be easily validated that all of the constraints used in those previous works can be implemented in the proposed method. Detailed comparison of the proposed method with the previous method are also presented in detail in this paper.

2. Related Works

Many methods using geometric constraints of a scene have been introduced. The work most closely related with that of the present paper was presented by Wilczkowiak et al. who suggested elegant linear formulation for camera calibration and 3D modeling using parallelepipeds in a scene [15]. For this method, at least six vertices of a parallelepiped must be visible in all views to obtain affine reconstruction. However, the primitive that can give full affine information is not only a parallelepiped. Parallelograms are most general primitives in a man-made scene, such as the architecture, but pairs of parallelograms do not always form

978-1-4244-3991-1/09/$25.00 ©2009 IEEE

two faces of a parallelepiped. It is also possible to obtain projective reconstruction linearly from at least one plane visible in all views [13]. However, to obtain metric reconstruction linearly, the method requires that the plane should be the plane at infinity determined from three orthogonal vanishing points as in [2]. Given calibration and rotation parameters, a linear estimation method using the information of metric geometry was also suggested in [3]. Quadratic constraints from coplanar parallelograms were discovered for camera calibration in [16]. These constraints can be converted to linear constraints with the metric information of the parallelograms.

3. The Infinity Homography from Parallelism and Coplanarity

3.1. Preliminaries and Motivations

In general, it is possible to estimate the infinite homography, $\mathbf{H}_\infty$, with at least four vanishing point correspondences in two views by solving the relevant linear equations of

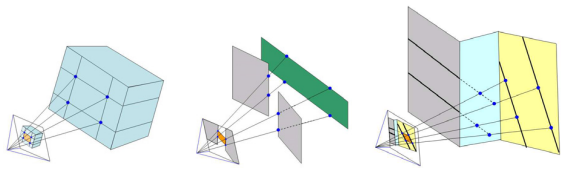
$$\mathbf{v}'_i \cong \mathbf{H}_\infty \mathbf{v}_i, \text{ for } i \in 1, \ldots, 4, \tag{1}$$

where $\mathbf{v}$ and $\mathbf{v}'$ are homogeneous coordinates of the corresponding vanishing points in two views and '$\cong$' indicates equality up to scale. The vanishing point can be estimated by the intersection of imaged lines parallel in 3D space. One of the vanishing points can be replaced by the epipole related to the two views, as it is mapped between the views by the infinite homography [5].

The two sets of parallel lines with different directions construct imaged parallelograms in camera images. If the two sets of parallel lines lie on the same supporting plane in 3D, the imaged parallelogram corresponds to an actual parallelogram existing on the plane in 3D (See Fig. 1(a)). This is common in typical man-made scenes. However, an imaged parallelogram consisting of two sets of parallel lines on a different supporting plane can exist, as in Fig. 1(b) and 1(c). In this case, the parallelogram does not actually exist in 3D. The proposed framework for computing $\mathbf{H}_\infty$ utilize the coplanarity constraint related to the former case.

3.2. Image Rectification for A Novel Framework

The rectification homographies $\mathbf{H}_{ri}$, for $i = 1, 2$, are defined here for two views. The rectification homography is defined so that it maps each imaged parallelogram to a rectangle that has edges parallel to the vertical and horizontal axes of the image plane (see Fig. 2). The position and the shape of the rectangle can be chosen arbitrarily. The rectification homography is allocated to each imaged parallelogram in views and is easily computed by using four vertex correspondences.



(a) (b) (c)

Figure 1. Examples of imaged parallelograms in camera images due to the two sets of parallel lines with different directions: Only (a) depicts the imaged parallelogram corresponding to an actual parallelogram existing on a plane in 3D.

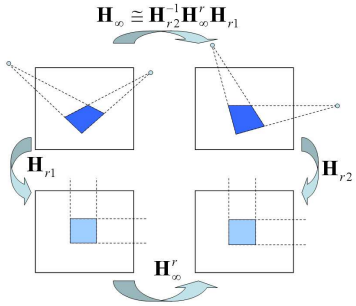


Figure 2. Relationship between the original $\mathbf{H}_\infty$ and the newly defined $\mathbf{H}_\infty^r$, which is the infinite homography between the rectified images.

The image transformed by rectification homography can be considered to be an image on a new camera image plane. A *rectified camera* is defined as a camera having this new image. This camera can be considered as a virtual camera acquired by rotating the original camera about its optical center and varying the intrinsic parameters. It is worthwhile to note that the image plane of a rectified camera is parallel to the related parallelogram in 3D.

In the images transformed using $\mathbf{H}_{ri}$, it is clear that the two vanishing points deduced from the transformed parallelogram, which is rectangle, are $[1, 0, 0]^T$ and $[0, 1, 0]^T$ in the homogenous coordinate system without any calculation because the parallel lines intersect in infinity (see Fig. 2).

Since the newly deduced vanishing points from the rectangle are also a transformation of the original vanishing points, it is possible to conclude that the infinite homography, $\mathbf{H}_\infty^r$, between the rectified cameras satisfies the following equations:

$$\lambda_x \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T = \mathbf{H}_\infty^r \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \tag{2}$$

and

$$\lambda_y \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T = \mathbf{H}_\infty^r \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T, \tag{3}$$

where λ_x and λ_y are arbitrary scale factors. From Eqs. (2)

and (3), $\mathbf{H}_\infty^r$ has the following form:

$$\mathbf{H}_\infty^r = \begin{bmatrix} \lambda_x & 0 & u \\ 0 & \lambda_y & v \\ 0 & 0 & w \end{bmatrix}, \quad (4)$$

where u, v, and w are arbitrary scalar values.

From Fig. 2, the relationship between the original $\mathbf{H}_\infty$ and the newly defined $\mathbf{H}_\infty^r$ is as follows:

$$\mathbf{H}_\infty \cong \mathbf{H}_{r2}^{-1}\mathbf{H}_\infty^r\mathbf{H}_{r1}. \quad (5)$$

This gives equations for the elements of $\mathbf{H}_\infty$ and $\mathbf{H}_\infty^r$.

As Eq. (5) can be obtained for each parallelogram (two sets of parallel lines), the number of equations constraining the variables is $9m+1$ (1 is for scale factor determination), where m is the number of parallelograms. Since the number of variables are 9 for $\mathbf{H}_\infty$ and $5m$ for $\mathbf{H}_\infty^{r_k}$ ($k = 1, \cdots, m$), $9m + 1 \geq 9 + 5m$ is required for the estimation to be possible. Thus, the required number of parallelograms is $m \geq 2$. It is worthwhile to note that this requirement is also identical for the method that uses vanishing point correspondences because two parallelogram can give four vanishing points. It can be concluded that, in the new framework, there is no change in the number of constraints. Accordingly, although Eq. (5) is used to compute $\mathbf{H}_\infty$, no advantage is given without using the additional constraint depicted in Section 3.1 .

3.3. An Additional Constraint from an Actual Parallelogram

This section verifies that one additional constraint applies to $\mathbf{H}_\infty^r$ when all lines included in the two sets of parallel lines related to the imaged parallelogram lie on the same plane in 3D.

First, the camera calibration matrix of a rectified camera is deduced. The camera calibration matrix is

$$\mathbf{K} = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where f_u and f_v denote the focal length expressed in horizontal and vertical pixel dimensions, respectively, s is the skew parameter, and (u_0, v_0) are the pixel coordinates of the principal point.

The *aspect ratio* of a parallelogram is defined as the ratio of the height to the length of the bottom side. Let the four vertices of a parallelogram in 3D be $[0,0,1]^T$, $[1,0,1]^T$, $[rcot\theta+1, r, 1]^T$, and $[rcot\theta, r, 1]^T$, where r is the aspect ratio of the parallelogram and θ is the angle between the two sides. $[0,0,1]^T$, $[1,0,1]^T$, $[1,a_i,1]^T$, and $[0,a_i,1]^T$, for $i = 1, 2$, are chosen as the four vertices of two rectangles in each view of two rectified cameras, where a_i, for $i = 1, 2$, are the aspect ratios of the two rectangles.

In fact, it will be shown that the variables r and θ disappear from the equation for the additional constraint, thus, the metric information of the parallelogram is not necessary.

The homography mapping between the vertices of the parallelogram and the rectangle of ith camera can then be computed as

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{h}_i^1 & \mathbf{h}_i^2 & \mathbf{h}_i^3 \end{bmatrix} \cong \begin{bmatrix} 1 & -cot\theta & 0 \\ 0 & a_i/r & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

The circular points of the plane containing the parallelogram are imaged on the rectified image to be $\mathbf{h}_i^1 \pm j\mathbf{h}_i^2$. Let $\boldsymbol{\omega}_{ri}$ be the image of the absolute conic (IAC) on the rectified image plane of the ith camera. It is then possible to obtain two equations that are linear in $\boldsymbol{\omega}_{ri}$ [8]

$$(\mathbf{h}_i^1 \pm j\mathbf{h}_i^2)^T\boldsymbol{\omega}_{ri}(\mathbf{h}_i^1 \pm j\mathbf{h}_i^2) = 0$$

or

$$\mathbf{h}_i^{1T}\boldsymbol{\omega}_{ri}\mathbf{h}_i^2 = 0 \quad \text{and} \quad \mathbf{h}_i^{1T}\boldsymbol{\omega}_{ri}\mathbf{h}_i^1 = \mathbf{h}_i^{2T}\boldsymbol{\omega}_{ri}\mathbf{h}_i^2 \ . \quad (7)$$

From Eq. (7), it can be seen that $\boldsymbol{\omega}_r$ has the form of

$$\boldsymbol{\omega}_{ri} \cong \begin{bmatrix} sin^2\theta & (r/a_i)sin\theta cos\theta & \alpha \\ (r/a_i)sin\theta cos\theta & r^2/a_i^2 & \beta \\ \alpha & \beta & \gamma \end{bmatrix},$$

where α, β, and γ are arbitrary scalar values.

If $\mathbf{K}_{ri}$ represents the camera calibration matrix of the ith rectified camera, then, because $\boldsymbol{\omega}_{ri} = \mathbf{K}_{ri}^{-T}\mathbf{K}_{ri}^{-1}$ [5], after some manipulations it is possible to obtain

$$\mathbf{K}_{ri} \cong \begin{bmatrix} sin\theta & -cos\theta & \alpha' \\ 0 & (a_i/r)sin\theta & \beta' \\ 0 & 0 & \gamma' \end{bmatrix}, \quad (8)$$

where α', β', and γ' are arbitrary scalar values.

As the circular points are fixed under a similarity transformation, the aforementioned selection of the vertices of the parallelogram and the rectangle does not lose generality.

Theorem 1 *The rotation matrix between two rectified cameras related to the same parallelogram is*

$$\mathbf{R}_I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \textit{ or } \mathbf{R}_{\bar{I}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proof: Without loss of generality, it is assumed that the plane supporting the parallelogram is on $Z = 0$ of the world coordinate system and that $[\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, \mathbf{t}]$ is the rotation and translation which relate the world coordinate system to the camera coordinate system. It can then be seen that

$$\mathbf{H}_i \cong \mathbf{K}_{ri}\begin{bmatrix} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{bmatrix}$$

or

$$\mu \mathbf{K}_{ri}^{-1} \mathbf{H}_i = \left[\begin{array}{ccc} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{array} \right], \quad (9)$$

where $\mathbf{H}_i$ corresponds to Eq. (6), $\mathbf{K}_{ri}$ corresponds to Eq. (8), and μ is an arbitrary scale factor. By substituting Eq. (6) and Eq. (8) into Eq. (9), the result is

$$\mu \left[\begin{array}{ccc} sin\theta & 0 & \alpha'' \\ 0 & sin\theta & \beta'' \\ 0 & 0 & \gamma'' \end{array} \right] = \left[\begin{array}{ccc} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{array} \right], \quad (10)$$

where α'', β'', and γ'' are arbitrary scalar values.

Considering that $\|\mathbf{r_1}\| = \|\mathbf{r_2}\| = \mathbf{1}$ and $\mathbf{r_3} = \mathbf{r_1} \times \mathbf{r_2}$, it is given that $[\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}] = \mathbf{R}_I$ or $\mathbf{R}_{\bar{I}}$. Since the rotation matrix between two rectified cameras is one of $\mathbf{R}_I^T \mathbf{R}_I$, $\mathbf{R}_I^T \mathbf{R}_{\bar{I}}$, $\mathbf{R}_{\bar{I}}^T \mathbf{R}_{\bar{I}}$, and $\mathbf{R}_{\bar{I}}^T \mathbf{R}_I$, it is $\mathbf{R}_I$ or $\mathbf{R}_{\bar{I}}$. ■

Theorem 2 *Assume that parallelograms imaged in two views are related to the same parallelogram which actually exists in 3D, and that the rectified parallelograms, which are rectangles, in each view have the aspect ratio of a_1 and a_2, respectively, then $\lambda_y = (a_2/a_1)\lambda_x$ in Eq. (4).*

Proof: Let $\mathbf{K}_{r1}$ and $\mathbf{K}_{r2}$ be the camera calibration matrices of the two rectified cameras. Assume $\mathbf{R_r}$ is the rotation matrix between two rectified cameras. From Theorem 1 and Eq. (8), $\mathbf{H}_\infty^r$ can be computed as follows [9]:

$$\mathbf{H}_\infty^r \cong \mathbf{K}_{r2} \mathbf{R}_r \mathbf{K}_{r1}^{-1} \cong \left[\begin{array}{ccc} 1 & 0 & u' \\ 0 & \frac{a_2}{a_1} & v' \\ 0 & 0 & w' \end{array} \right],$$

where u', v', and w' are arbitrary scalar values. ■

As mentioned above, it is worthwhile to note that $\mathbf{H}_\infty^r$ is independent of the angle between the two sides of the parallelogram, θ, and the aspect ratio of the parallelogram, r. From Theorem 2, if the two cameras observe the same parallelogram which actually exists in 3D, the additional and linear constraint of $\lambda_y = (a_2/a_1)\lambda_x$ can be inserted in Eq. (5) because the aspect ratios, a_1 and a_2, are values that can be chosen arbitrarily. If a parallelogram does not in fact exist in 3D, it is not possible to assume that r is identical for each camera due to the parallax of the two cameras (refer to Fig. 1(b) and 1(c)); the constraint cannot be used.

Assume that m parallelograms are viewed with two cameras and that $m_r (\leq m)$ is the number of actual parallelograms. From Eq. (5) and the additional constraint, the infinite homography, $\mathbf{H}_\infty$, can be computed by solving the set of homogeneous equations

$$\mathbf{H}_\infty \cong {\mathbf{H}_{r2}^k}^{-1} \mathbf{H}_\infty^{r_k} \mathbf{H}_{r1}^k, \text{ for } k = 1, \ldots, m \quad (11)$$

or

$$\mathbf{A}[H_{11}, H_{12}, \ldots, H_{33}, \lambda_x^1, (\lambda_y^1), u^1, v^1, w^1, \ldots, \lambda_x^m, (\lambda_y^m), u^m, v^m, w^m]^T = \mathbf{0}, \quad (12)$$

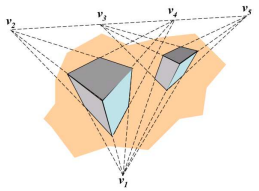


Figure 3. Examples of the position of vanishing points. Four vanishing points are collinear.

where $\mathbf{A}$ is a $9m \times (9+5m-m_r)$ matrix composed of the element of the rectification homographies, $\mathbf{H}_{r1}^k$ and $\mathbf{H}_{r2}^k$, and H_{ij} is the (i, j) element of $\mathbf{H}_\infty$. The parentheses in Eq. (12) indicate that if kth parallelogram is an actual parallelogram, the variable, λ_y^k, is not necessary due to the additional constraint.

3.4. Comparison with Related Works

In general, four vanishing points are required to obtain the infinite homography or three vanishing points together with the fundamental matrix. However, it is difficult to acquire more than three vanishing points giving independent constraints in general man-made scenes as in Fig. 3 and to obtain many point correspondences between all view pairs for the fundamental matrix estimation. Even in that case, the proposed method can obtain the infinite homography due to the additional constrains arising from the actual parallelograms. A degenerate case of the proposed method occurs only when all parallelograms are on planes parallel to each other. Moreover, even if there are four vanishing points, the additional constraints can improve the accuracy of camera calibration when metric constraints are not sufficient. This issue is explored further in Section 6.1.

However, it can be seen that these results also can be acquired from the method of [15] because parallelepiped gives only three vanishing points. This is due to the fact that coplanarity constraints are merged in the canonical projection matrix suggested in [15]. However, as mentioned previously, the proposed method can deal with the general position of parallelograms in practical situations due to the additional constraints and, consequently, is free from the restriction that the parallelograms should be the faces of a parallelepiped.

Additional quadric constraints can also be derived from parallelograms and used to calibrate camera parameters directly [16]. However, to make overall reconstruction approach linear, the constraints need metric information to be converted to linear ones, such as rectangle and diamond. In fact, the linear constraints derived like this are implemented in Section 5 from equal length and orthogonality constraints after all.

4. Reconstruction up to Affine Transformation

Once the infinite homography, $\mathbf{H}^{0i}$, between one reference view and ith view are computed, the camera matrices of an affine reconstruction are $\mathbf{P}_0 = [\mathbf{I}|\mathbf{0}]$ and $\mathbf{P}_i = [\mathbf{H}_\infty^{0i}|\mathbf{t}_i]$ [5]. If ith view do not have common images of two or more parallelogrmas with reference view, some manipulations may be needed using intermediate view. If the infinite homographies of $\mathbf{H}_\infty^{ji}$ and $\mathbf{H}_\infty^{0j}$ are given, $\mathbf{H}_\infty^{0i}$ can be computed as follows:

$$\mathbf{H}_\infty^{0i} = \mathbf{H}_\infty^{ji}\mathbf{H}_\infty^{0j}.$$

The variables remaining to be solved are camera motion $\mathbf{t}$ and 3D points $\tilde{\mathbf{X}}$. The equation for point projection is

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \mathbf{P}\begin{bmatrix} \tilde{\mathbf{X}} \\ 1 \end{bmatrix} \cong \begin{bmatrix} \mathbf{h}_1^T & t_1 \\ \mathbf{h}_2^T & t_2 \\ \mathbf{h}_3^T & t_3 \end{bmatrix}\begin{bmatrix} \tilde{\mathbf{X}} \\ 1 \end{bmatrix}. \quad (13)$$

By taking vector product of the two sides of Eq. (13), two independent homogeneous equations

$$\begin{bmatrix} x\mathbf{h}_3^T - \mathbf{h}_1 & -1 & 0 & x \\ y\mathbf{h}_3^T - \mathbf{h}_2 & 0 & -1 & y \end{bmatrix}\begin{bmatrix} \tilde{\mathbf{X}} \\ \mathbf{t} \end{bmatrix} = \mathbf{0} \quad (14)$$

are obtained. Thus, $2nm$ set of equations in $3n + 3(m-1)$ unknowns is generated with m views and n 3D points. It is worthwhile to note that $\mathbf{t}$ of the reference view $\mathbf{P}_0$ is assumed to be $\mathbf{0}$ without loss of generality. These equations can be solved linearly to obtain $\mathbf{t}$ and $\tilde{\mathbf{X}}$ of an affine reconstruction. In this formulation, all points need not to be visible in all views.

4.1. Parameter Reduction Using Affine Invariant

If there are parallelograms or parallelepipeds in views, the number of variables can be reduced with geometric constraints. In affine reconstruction, all corner points of parallelograms or parallelepipeds are represented by one reference point and two vectors or three vectors. These are common properties with metric reconstruction. Moreover, the coordinate of vanishing point in first view are considered as vector parallel to the lines corresponding to the vanishing point when using the above affine reconstruction formulation. Representing the points on the surfaces and edges of parallelograms and parallelepipeds with weighted sum of these vectors, coplanarity and colinearity constraints and constraints of length ratios between line segments parallel to each other can be satisfied, which is affine invariant.

If $\mathbf{D}$ is column vector composed of above vectors ,scalar weights for these vectors, and other 3D points, $\tilde{\mathbf{X}}$ can be represented by

$$\tilde{\mathbf{X}} = \mathbf{\Sigma}\mathbf{D}, \quad (15)$$

where $\mathbf{\Sigma}$ contains the geometric constraints described above. Then, we must solve the following equations,

$$\begin{bmatrix} (x\mathbf{h}_3^T - \mathbf{h}_1)\mathbf{\Sigma} & -1 & 0 & x \\ (y\mathbf{h}_3^T - \mathbf{h}_2)\mathbf{\Sigma} & 0 & -1 & y \end{bmatrix}\begin{bmatrix} \mathbf{D} \\ \mathbf{t} \end{bmatrix} = \mathbf{0} \quad (16)$$

and the reconstruction results are obtained by Eq. (15). Due to this parameterization, the reconstructed vertices constitute exactly parallelograms and parallelepipeds and the coplanarity and colinearity constraints for the points on the surfaces and edges of parallelograms and parallelepipeds are exactly satisfied.

4.2. Comments on Linear Reconstruction Formula

The linear reconstruction approach described above is similar to those method presented in [13] and [3]. The camera having zero skew and known aspect ratio or static internal parameters is assumed in [13]. The problem of the method to calibrate internal parameter directly from orthogonal vanishing points as in [13] and [2] is a degenerate case where one or two of the vanishing points are at infinity. However, the restriction does not occur in estimation of the infinity homography and affine reconstruction from the information of parallelism. Moreover, the orthogonal information can be used as well in the process of the metric upgrade described in the next section. The structure parameter reduction is similar to [6] presenting single view based approach and is also suggested in the linear step in [3]. However, in [3], given that the calibration and rotation parameters are estimated previously by other calibration methods, the parameterization using the scene geometry is performed in metric space not in affine space. It is also straightforward that the symmetry and regularity information used in [3] is all implemented in the overall process of the proposed approach by using the constraints of length ratios in Section 4.1 and orthogonality in 5.1. This means that, in the proposed method, the full geometric information contributes to both camera calibration and reconstruction because the calibration results are obtained at the last step.

5. Upgrade to Metric Reconstruction

Let $\mathbf{H}_E^A$ be the projective transformation from metric to affine space and has the following form

$$\mathbf{H}_E^A = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}. \quad (17)$$

To convert affine reconstructions obtained in the previous section to metric ones, we should find constraints on $\mathbf{A}$. This is equivalent to obtain constraints on the absolute conic $\mathbf{\Omega}_\infty^A$ in affine space. $\mathbf{A}$ can be obtained from $\mathbf{\Omega}_\infty^A (= \mathbf{A}^{-T}\mathbf{A}^{-1})$ by Cholesky factorization.

5.1. Constraints from Scene Geometry

Assume that there are four points, $\tilde{\mathbf{X}}_{Ei}$, for $i = 1, \cdots, 4$, in metric space and two vectors, $\mathbf{d}_{E1} = \tilde{\mathbf{X}}_{E2} - \tilde{\mathbf{X}}_{E1}$ and $\mathbf{d}_{E2} = \tilde{\mathbf{X}}_{E4} - \tilde{\mathbf{X}}_{E3}$. If four points in affine space corresponding to the above four points are $\tilde{\mathbf{X}}_{Ai}$, for $i = 1, \cdots, 4$, and $\mathbf{d}_{A1} = \tilde{\mathbf{X}}_{A2} - \tilde{\mathbf{X}}_{A1}$ and $\mathbf{d}_{A2} = \tilde{\mathbf{X}}_{A4} - \tilde{\mathbf{X}}_{A3}$, then, $\mathbf{d}_{E1} = \mathbf{A}^{-1}\mathbf{d}_{A1}$ and $\mathbf{d}_{E2} = \mathbf{A}^{-1}\mathbf{d}_{A2}$. Assume that $\mathbf{d}_{E1}$ and $\mathbf{d}_{E2}$ are orthogonal to each other in metric space. Then, since $\mathbf{d}_{E1}^T\mathbf{d}_{E2} = 0$, we can obtain following constraint

$$\mathbf{d}_{A1}^T\mathbf{\Omega}_\infty^A\mathbf{d}_{A2} = 0. \tag{18}$$

This equation is similar to the equation describing the relation between the image of absolute conic (IAC) and orthogonal vanishing points. However, Eq. (18) needs only the reconstructed 3D points and do not need particular points such as vanishing points.

If we know the ratio of $\mathbf{d}_{E1}$ to $\mathbf{d}_{E2}$ as r, then,

$$\mathbf{d}_{E1}^T\mathbf{d}_{E1}/\mathbf{d}_{E2}^T\mathbf{d}_{E2} = r^2$$

or

$$\mathbf{d}_{A1}^T\mathbf{\Omega}_\infty^A\mathbf{d}_{A1} = r^2\mathbf{d}_{A2}^T\mathbf{\Omega}_\infty^A\mathbf{d}_{A2}. \tag{19}$$

Moreover, if we know additionally the angle between $\mathbf{d}_{E1}$ and $\mathbf{d}_{E2}$ as θ, then,

$$cos\theta = \frac{\mathbf{d}_{E1}^T\mathbf{d}_{E2}}{(\mathbf{d}_{E1}^T\mathbf{d}_{E1})^{1/2}(\mathbf{d}_{E2}^T\mathbf{d}_{E2})^{1/2}}$$

or

$$\mathbf{d}_{A1}^T\mathbf{\Omega}_\infty^A\mathbf{d}_{A2} = rcos\theta\mathbf{d}_{A2}^T\mathbf{\Omega}_\infty^A\mathbf{d}_{A2}. \tag{20}$$

It is worthwhile to note that all equations derived in this section can be solved linearly.

5.2. Incorporation of Self-calibration Constraints

The absolute dual quadric (ADQ), $\mathbf{Q}_\infty^{*A}$, in affine space can be computed from $\mathbf{H}_E^A\mathbf{Q}_\infty^{*E}{\mathbf{H}_E^A}^T$ [5], where $\mathbf{Q}_\infty^{*E}$ is ADQ in metric space, and has the following form

$$\mathbf{Q}_\infty^{*A} = \begin{bmatrix} (\mathbf{\Omega}_\infty^A)^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}. \tag{21}$$

The IAC, ω_i, in ith view can be related to $\mathbf{Q}_\infty^{*A}$ as follows [5]:

$$\omega_i = (\mathbf{P}_i\mathbf{Q}_\infty^{*A}{\mathbf{P}_i}^T)^{-1} = {\mathbf{H}_\infty^{0i}}^{-T}\mathbf{\Omega}_\infty^A{\mathbf{H}_\infty^{0i}}^{-1}. \tag{22}$$

If the camera is static, we can set all $\{\omega_i\}$ to be ω in Eq. (22) and, then, ω can be eliminated using the relation $\omega = \mathbf{\Omega}_\infty^A$ in reference view. These results in linear equations on $\mathbf{\Omega}_\infty^A$ provided $\mathbf{H}_\infty^{0i}$ is normalized as det($\mathbf{H}_\infty^{0i}$)=1.

Knowing that pixel is square, and principal point is at origin, and aspect ratio is known as r, the following linear equations on $\mathbf{\Omega}_\infty^A$ can be written from Eq. (22):

$$\begin{cases} ({\mathbf{H}_\infty^{0i}}^{-T}\mathbf{\Omega}_\infty^A{\mathbf{H}_\infty^{0i}}^{-1})_{12} = 0 \\ ({\mathbf{H}_\infty^{0i}}^{-T}\mathbf{\Omega}_\infty^A{\mathbf{H}_\infty^{0i}}^{-1})_{13} = ({\mathbf{H}_\infty^{0i}}^{-T}\mathbf{\Omega}_\infty^A{\mathbf{H}_\infty^{0i}}^{-1})_{23} = 0 \\ r^2({\mathbf{H}_\infty^{0i}}^{-T}\mathbf{\Omega}_\infty^A{\mathbf{H}_\infty^{0i}}^{-1})_{11} = ({\mathbf{H}_\infty^{0i}}^{-T}\mathbf{\Omega}_\infty^A{\mathbf{H}_\infty^{0i}}^{-1})_{22}. \end{cases}$$

5.3. Comparison with Related Works

The constraints in Section 5.1 and 5.2 can be derived by basic knowledge of projective geometry. These constraints are somewhat equivalent to the constraints in [15] that is extracted by the prior knowledge about intrinsic parameters of cameras and parallelepipeds. The difference is that the 3D points used to derive the constraints are not only the vertices of parallelograms or parallelepipeds but also any 3D points in a scene in the proposed approach.

6. Experimental Results

Various experiments with real images captured for indoor and outdoor scenes were also performed to test the algorithms.

6.1. Tower Scene

The resolution of the images was 1024×768. Fig. 4 shows the images used in this experiment. The lines on the planes were extracted and the vertices of the parallelograms were extracted from the intersections of the lines. The two parallelograms denoted in Fig. 4(a) with the white dotted lines were used as the input for the algorithms. The prior information used for the algorithm was camera having a static intrinsic parameters. Fig. 4(d) shows the reconstructed model and the camera pose. Rendered new views of the reconstructed model are shown in Fig. 4(e) and 4(f). Since the ground truth or any references for the reconstruction results are not available in this experiment, the accuracy of the reconstruction for the known geometry that is nevertheless not used in the algorithm is a useful measure of the performances of the algorithms. The line a and the plane consisting of the two lines b and c depicted in Fig. 4(c) were reconstructed. It was known that the angle between the line a and the normal to the plane was zero. In this experiment, there were four vanishing points and the method using vanishing point correspondences also can be used. The measures of this angle were 2.22^o and 16.81^o using each method. From various experiments not presented here due to the lack of space, it was found that the proposed method gave more accurate calibration results owing to the additional constraints even in the case of four vanishing points when the static camera constraint was only used. If the geometric constraints were sufficient in this case, the results from the two methods were comparable.

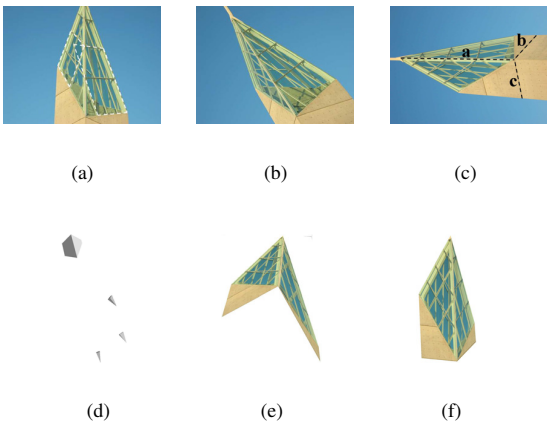


Figure 4. Three images used in the *Tower Scene* experiment and reconstructed model and camera pose.

6.2. Plant Scene

The resolution of the images was 1024×768 and the camera parameters were not static while the images were captured. Four captured images are shown in Fig. 5. *Image 0* corresponds to the reference camera. The lines on the faces of the building were extracted and the points to be reconstructed were extracted from the intersections of the lines. The infinite homographies $\mathbf{H}_{\infty}^{01}$, $\mathbf{H}_{\infty}^{12}$, and $\mathbf{H}_{\infty}^{23}$ are computed in this experiments. $\mathbf{H}_{\infty}^{01}$ can be obtained with the two parallelograms corresponding to the frontal and right face B of the building. However, it looks like that one parallelogram is insufficient between *Image 1* and *Image 2* because at least two parallelograms not parallel to each other are required to obtain sufficient constraints. In this case, it can be considered that the right and left faces (B and C) of the building are same parallelograms. This is due to the fact that the proposed method computing the infinite homogrphy is independent to the translational motion of the parallelograms. So, face A, B, and C are used to compute $\mathbf{H}_{\infty}^{12}$. Thus, in contrast to the method described in [15], since the proposed method can use parallelograms and partially overlapped images, it is more flexible in use. For $\mathbf{H}_{\infty}^{23}$, face A and C are used. Metric constraints used in this experiment were: orthogonality of the edge of the building, right angle between line segment a and b, and length ratio of c to d. Fig. 6 shows the reconstructed model and the camera pose in new view positions.

6.3. Scene of *Bank of China*

The images of *Bank of China* were gathered from the internet(see Fig. 7). The resolution of the images was various from 419×783 to 1536×2048. $\mathbf{H}_{\infty}^{01}$, $\mathbf{H}_{\infty}^{02}$, $\mathbf{H}_{\infty}^{23}$, and $\mathbf{H}_{\infty}^{24}$ are computed using $\{A, C\}$, $\{A, B, D\}$, $\{D, E\}$, and $\{A, D, F, G\}$, respectively. It is worthwhile to note that

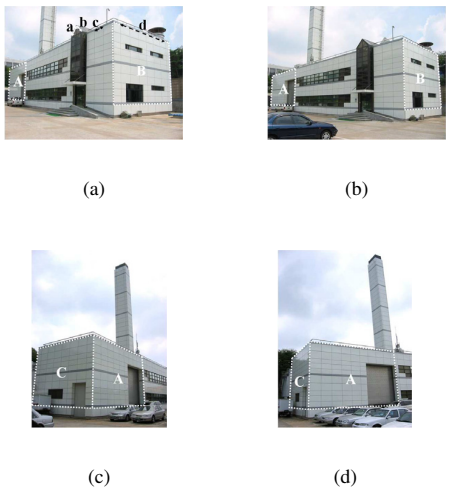


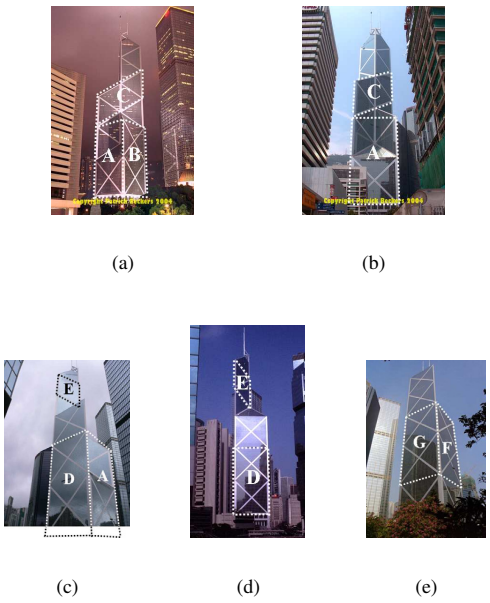
Figure 5. Four captured images for *Plant Scene* experiment. *Image* (a) *0*, (b) *1*, (c) *2*, (d) *3*

Figure 6. Reconstructed model and camera pose for *Plant Scene* experiment.

there are no parallelepipeds of which at least six vertices are seen across *Image 0* and *Image 2* but we can find common parallelograms. It is also the same for the pair of *Image 2* and *Image 3*. $\{A, G\}$ and $\{D, F\}$ can be considered as the same parallelograms respectively in the estimation of $\mathbf{H}_{\infty}^{24}$ as explained in Section 6.2. Metric constraints used in this experiment were: orthogonality of the edge of the building and orthogonality of crossing diagonal line. Reconstructed model and the camera pose are shown in Fig. 8.

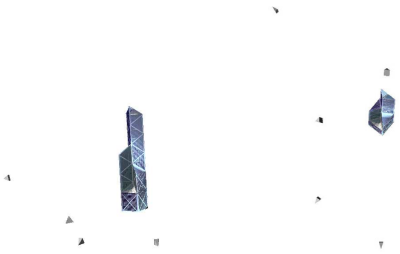
7. Conclusions

In this paper, a novel framework was introduced to reconstruct a 3D model using uncalibrated views and geometric information of a model. The proposed method is a stratified and linear method. It was shown that all of the constraints derived in the previous works can be implemented in the proposed method. The study on the situations that can not be dealt with by the previous approaches was presented and it was shown that the proposed method is more flexible in that the cases can be handled by



(a) (b) (c) (d) (e)

Figure 7. Five images for the experiment of the scene of *Bank of China*. *Image* (a) *0*, (b) *1*, (c) *2*, (d) *3*, (e) *4*



(a) (b)

Figure 8. Reconstructed model and camera pose for experiment of *Bank of China*.

the proposed method.

Acknowledgments

This work was supported by the IT R&D program of MKE/IITA and MCST/KOCCA. [2007-S-051-02, Software Development for Digital Creature].

References

[1] L. Agapito, E. Hayman, and I. Reid. Self-calibration of rotating and zooming cameras. *International Journal of Computer Vision*, 45(2):107–127, Nov. 2001.

[2] R. Cipolla, T. Drummond, and D. P. Robertson. Camera calibration from vanishing points in images of architectural scenes. In *Proc. British Machine Vision Conferece*, pages 382–391, Nottingham, England, Sept. 1999.

[3] E. Grossmann and J. Santos-Victor. Least-squares 3D reconstruction from one or more views and geometric clues. *Computer Vision and Image Understanding*, 99(2):151–174, Aug. 2005.

[4] P. Hammarstedt, F. Kahl, and A. Heyden. Affine reconstruction from translational motion under various autocalibration constraints. *Journal of Mathematical Imaging and Vision*, 24(2):245–257, Mar. 2006.

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

[6] D. Jelinek and C. J. Taylor. Reconstruction of linearly parameterized models from single images with a camera of unknown focal length. *IEEE Trans. Pattern Anal. Machine Intell.*, 23(7):767–773, July 2001.

[7] J.-S. Kim and I. S. Kwon. Estimating intrinsic parameters of cameras using two arbitrary rectangles. In *Proc. International Conference on Pattern Recognition*, Hong Kong, Aug. 2006.

[8] D. Liebowitz and A. Zisserman. Combining scene and autocalibration constraints. In *Proc. IEEE International Conference on Computer Vision*, pages 293–300, Kerkyra, Greece, Sept. 1999.

[9] Q.-T. Luong and T. Viéville. Canonical representations for the geometries of multiple perspective views. *Computer Vision and Image Understanding*, 64(2):193–229, Sept. 1996.

[10] T. Moons, L. V. Gool, M. Proesmans, and E. Pauwels. Affine reconstruction from perspective image pairs with a relative object-camera translation in between. *IEEE Trans. Pattern Anal. Machine Intell.*, 18(1):77–83, Jan. 1996.

[11] M. Pollefeys. Visual modeling with a hand-held camera. *International Journal of Computer Vision*, 59(3):207–232, Oct. 2004.

[12] M. Pollefeys and L. V. Gool. Stratified self-calibration with the modulus constraint. *IEEE Trans. Pattern Anal. Machine Intell.*, 21(8):707–724, Aug. 1999.

[13] C. Rother and S. Carlsson. Linear multi view reconstruction and camera recovery using a reference plane. *International Journal of Computer Vision*, 49(2-3):117–141, Sept. 2002.

[14] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Trans. Robot. Automat.*, 3(4):323–344, Aug. 1987.

[15] M. Wilczkowiak, P. Sturm, and E. Boyer. Using geometric constraints through parallelepipeds for calibration and 3D modelling. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(2):194–207, Feb. 2005.

[16] F. C. Wu, F. Q. Duan, and Z. Y. Hu. An affine invariant of parallelograms and its application to camera calibration and 3D reconstruction. In *Proc. European Conference on Computer Vision*, pages 191–204, May 2006.

[17] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Machine Intell.*, 22(11):1330–1334, Nov. 2000.