A New Linear Method for Euclidean Motion/Structure from Three Calibrated Affine Views

Long QUAN
CNRS-GRAVIR-INRIA
ZIRST – 655 avenue de l'Europe
38330 Montbonnot, France

Yuichi OHTA
University of Tsukuba
Institute of Information Sciences
Tsukuba, Ibaraki, 305-8573, Japan

Abstract

We introduce a unified framework for developing matching constraints of multiple affine views and rederive 2-view (affine epipolar geometry) and 3-view (affine image transfer) constraints within this framwork. We then describe a new linear method for Euclidean motion and structure from 3 calibrated affine images, based on insight into the particular structure of these multiple-view constraints.

Compared with the existing linear method of Huang and Lee [7], the new method uses different and more appropriate constraints. It has no failure mode of the Euclidean factorisation method of Tomasi and Kanade [18]. We demonstrate the method on real image sequences.

1 Introduction

Motion/structure from orthographic or weak perspective views is a very old and popular topic. It is well known that at least 4 non-planar points over 3 orthographic or weak perspective views are sufficient to uniquely determine motion/structure up to a reflection about the image plane [20, 7, 8]. Many algorithms have been published for this problem: the linear methods of Huang and Lee [7, 9], non-linear algebric methods of Koenderink and Van Doorn [8, 2] and non-linear numerical method of Shapiro *et al.* [16]. A good review of the different methods can be found in [16].

In this paper, we introduce a unified framework for developing matching constraints of multiple affine views. In particular, 2-view and 3-view constraints will be derived, and all existing methods could be recast into this framework. Our key observation is that classical linear methods for metric motion/structure from 3 orthographic or weak perspective views were heavily based on affine epipolar geometry and did not use the full set of 3 image constraints, thus leading to over-parameterisation and inconsitent motion recovery. Based on insight into the particular structure of these constraints, we will propose a linear algorithm that uses 9 linear parameters to encode the 8 Euclidean motion parameters of 3 weak perspective views.

This work was initially motivated by the application of novel image synthesis from example images, as Euclidean reconstruction from a minimal number of images is required here [10, 23, 24].

The paper is organised as follows. In Section 2, we review the affine camera model. Then, we introduce a unified framwork for studying the geometric constraints among multiple affine images in Section 3. The linear method for Euclidean motion/structure from 3 calibrated affine images is developed in Section 4. Experimental results are presented in Section 5 and a short conclusion is given in Section 6

Throughout the paper, matrices are denoted in upper case boldface, vectors in lower case boldface and scalars are either in lower case or lower case Greek.

2 Review of the affine camera model

For a restricted class of camera models, by setting the third row of the perspective camera $\mathbf{P}_{3\times 4}$ to $(0,0,0,\lambda)$, we obtain the affine camera initially introduced by Mundy and Zisserman in [11, 1]

$$\mathbf{A}_{3\times 4} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{34} \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{2\times 3} & \mathbf{t}_{3\times 1} \\ \mathbf{0}_{1\times 3} & \end{pmatrix}.$$

The affine camera $\mathbf{A}_{3\times 4}$ subsumes the orthographic, weak perspective and para-perspective. For more detailed relations and applications, one can refer to [16, 14].

Finite points in affine spaces $\mathcal{R}^n$, are naturally embedded into $\mathcal{P}^n$ by the mapping $\mathbf{u}_a \mapsto \mathbf{u} = (\mathbf{u}_a, 1)^T$ and $\mathbf{x}_a \mapsto \mathbf{x} = (\mathbf{x}_a, 1)^T$. We have therefore $\mathbf{u}_a = \mathbf{M}_{2\times 3}\mathbf{x}_a + \mathbf{t}_0$, where $\mathbf{t}_0 = (t_1/t_3, t_2/t_3)^T = (p_{14}/p_{34}, p_{24}/p_{34})^T$. If we further use relative coordinates of the points with respect to a given reference point (for instance, the centroid of a set of points), the vector $\mathbf{t}_0$ is cancelled, and we obtain the following linear

mapping between relative space points and relative image points:

$$\Delta\mathbf{u} = \mathbf{M}_{2\times 3}\Delta\mathbf{x}. \tag{1}$$

Equation (1) is the basic projection equation for points in an affine camera when relative coordinates are used. The reference point to determine the relative coordinates determines uniquely the translational component of the affine projection matrix. Throughout this paper, the reference point is always taken to be the centroid in each image.

3 Unifying 2-view and 3-view geometry of points

For projective cameras, the geometric constraints among multiple projective views have been thoroughly studied in [19, 17, 6, 5]. There has been no similar effort for affine camera case, although the geometric constraints among affine views are well known.

For notational simplicity, rewrite Equation (1) as $\mathbf{u} = \mathbf{M}_{2\times 3}\mathbf{x}$.

We can now examine the matching constraints between multiple views of the same point. Let the three views of the same point $\mathbf{x}$ be given as follows:

$$\begin{cases} \mathbf{u} &= \mathbf{M}\mathbf{x}, \\ \mathbf{u}' &= \mathbf{M}'\mathbf{x}, \\ \mathbf{u}'' &= \mathbf{M}''\mathbf{x}. \end{cases} \tag{2}$$

These can be rewritten together in matrix form as

$$\begin{pmatrix} \mathbf{M} & \mathbf{u} \\ \mathbf{M}' & \mathbf{u}' \\ \mathbf{M}'' & \mathbf{u}'' \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \lambda \end{pmatrix} = 0, \tag{3}$$

where $\lambda \neq 0$ encodes the (unrecoverable) global scale factor of the reconstruction.

As the vector $(\mathbf{x}, \lambda)^T$ can not be zero, the rank of the coefficient matrix is at most 3, so all of its 4×4 minors vanish. There are $C_6^4 = 15 = 3+4+4+4$ such minors, which can be divided into two types:

- 2-view constraints involving only two views with two rows from each view,
- 3-view constraints involving all three views with two rows from one view and one from each of the others.

There are three 2-view and three sets of four 3-view constraints. Among the 3 sets of 3-view constraints, only one of them is independent due to the symmetry.

Each expansion of these 4×4 minors is linear in the image coordinates $\mathbf{u}$, $\mathbf{u}'$ and $\mathbf{u}''$ with the coefficients t_i coming from the 3×3 minors of the following 6×3 joint projection matrix: $(\mathbf{M}^T, \mathbf{M}'^T, \mathbf{M}''^T)^T = (1\ 2\ 1'\ 2'\ 1''\ 2'')^T$.

There are in total $C_6^3 = 20 = 8 + 4 + 8$ such minors, as we will see later that 4 of the 20 minors are common to the 2-view and 3-view constraints. All these minors provide a linear coordinate system to span the joint projection matrix. The constraints for more than 3 views will be briefly discussed in Section 6.

3.1 Two-view constraints

There are three 2-view constraints corresponding to the 3 pairs of the 3 views, namely the vanishing of the determinants $[121'2']$, $[121''2'']$ and $[1'2'1''2'']$:

$$\begin{aligned} t_{13}u &+ t_{14}v &+ t_{10}u' &+ t_9\,v' &= 0, \\ t_{15}u' &+ t_{16}v' &+ t_{17}u'' &+ t_{18}v'' &= 0, \\ t_{19}u &+ t_{20}v &+ t_{12}u'' &+ t_{11}v'' &= 0. \end{aligned}$$

These are the affine epipolar geometry. The set of $3 \times 4 = 12$ coefficients t_i for $i = 9, \ldots, 20$ are 12 of the 20 minors of the joint projection matrix.

Each point correspondence from two images gives one homogeneous linear equation, taking into account of the reference point for relative coordinates, 4 points are sufficient for uniquely determining the affine epipolar geometry.

The affine epipolar geometry equation was introduced in [7] for orthography, and later in [9] for weak perspective and also in [16]. Shapiro *et al.* nicely related the affine epipolar geometry with Koenderink and Van Doorn's rotation parameterisation [8]. But Koenderink's method is equivalent to Lee and Huang's [9].

3.2 Three-view constraints

There are four 3-view constraints from the vanishing of the determinants $[121'1'']$, $[122'1'']$, $[121'2'']$ and $[122'2'']$. By careful inspection of the minors (for example, using the computer algebra tool Maple), we have:

$$\begin{aligned} t_4\,u &+ t_8\,v &+ t_{11}u' &+ t_9\,u'' &= 0, \\ t_2\,u &+ t_6\,v &+ t_{11}v' &+ t_{10}u'' &= 0, \\ t_3\,u &+ t_7\,v &+ t_{12}u' &+ t_9\,v'' &= 0, \\ t_1\,u &+ t_5\,v &+ t_{12}v' &+ t_{10}v'' &= 0. \end{aligned}$$

Among 12 minors, 8 of them are new (t_1 to t_8) and 4 of them are common with 2-view constraints (t_9 to t_{12}).

These are the transfer equations over three views [21]. Any orthographic view of a point set can be expressed as a linear combination of two other views if this point set undergoes a linear transformation in space. This has been extensively used in object recognition.

Since each point correspondence gives 4 linearly independent 3-view constraints, 4 points give $(4-1) \times 4 = 12$ linear equations for solving these minors.

The appearance of common minors between 2-view and 3-view constraints is not accidental, as we have 3+4=7 constraints, each of them has 4 coefficients, that amounts to $4 \times 7 = 28$. As there are only 20 minors, so 8 of them should appear at least more than once.

4 Euclidean motion/structure from 3 calibrated affine views

So far, the linear estimates of the 2-view and 3-view constraints yield, directly but implicitly, the affine motion/structure. To get Euclidean or more exactly similarity motion/structure, we need at least 3 calibrated affine images. Here we use the unified formulation introduced in [14] for calibrated affine cameras, thus the method developed will be valid for all calibrated orthographic, weak-perspective and para-perspective models.

Each projection matrix $\mathbf{M}_{2\times 3}$ can be decomposed [14] into $\mathbf{M} = s\mathbf{K}\mathbf{R}$, where s is a scaling factor of the whole image, $\mathbf{K}$ is the intrinsic parameter matrix (for instance, $\mathbf{K} = \begin{pmatrix} \xi & 0 \\ 0 & 1 \end{pmatrix}$, where ξ is the aspect ratio for the weak-perspective case), and $\mathbf{R}_{2\times 3}$ represents 2 rows of a 3D rotation matrix. As we are assuming calibrated cameras, the intrinsic parameter matrix $\mathbf{K}$ is known and its inverse can be directly applied to the image points so that its effect is removed completely. So the projection matrix $\mathbf{M}$ for normalised image coordinates becomes $\mathbf{M} = s\mathbf{R}$, i.e. a scaled 2×3 rotation matrix. There are in total $8 = 2 \times 3 + 2$ Euclidean parameters for a set of 3 views: two relative 3D rotations $\mathbf{R}$ and $\mathbf{G}$ each having 3 d.o.f. and 2 relative scale factors s and s' between (say) the first view and the remaining ones.

Any linear algorithm will consist of first estimating linearly the minors using multiple view constraints, then extract the 8 Euclidean parameters from these minors by identifying the projection matrices $\mathbf{M}$ and $\mathbf{M}'$ with the scaled rotation matrices $s\mathbf{R}$ and $s'\mathbf{G}$.

Combining 2-view and 3-view constraints We should keep in mind that although all 7 constraints are linearly independent, only $3 = (6-3) \times (4-3)$ of them are algebraically independent due to Grassmanian relations. There are in total 20 homogeneous coefficients—minors of the joint projection matrix. How to choose the most appropriate constraints is of primary importance. The selected constraints should be algebraically independent and contain as few coefficients as possible.

Taking only the three 2-view constraints is a poor choice since the third one is partially dependent on the first two by the composition rule on the rigid motions and each one is completely separate from the others. Taking only the 3-view constraints leads to a complicated algebraic manipulation for the extraction of Euclidean parameters. Hence we combine 2-view and 3-view constraints.

The key observation is that there are common coefficients between 2-view and 3-view constraints: two of the three 2-view constraints share 4 minors t_9, t_{10}, t_{11} and t_{12} with the 3-view constraints. This allows us to use the following combination: two 2-view constraints plus one of the 3-view constraints.

$$\begin{cases} t_4\, u & +t_8\, v & +t_{11}u' & +t_9\, u'' & = & 0, \\ t_{13}u & +t_{14}v & +t_{10}u' & +t_9\, v' & = & 0, \\ t_{15}u & +t_{16}v & +t_{12}u'' & +t_{11}v'' & = & 0. \end{cases}$$

These 10 unknown minors can be solved as a single homogeneous vector under the constraint $||\mathbf{t}_{10}|| = 1$:

$$\begin{pmatrix} u & v & u'' & 0 & u' & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & v' & u' & 0 & 0 & u & v & 0 & 0 \\ 0 & 0 & 0 & 0 & v'' & u'' & 0 & 0 & u & v \end{pmatrix} \mathbf{t} = 0.$$

Any ratio t_i/t_j of the minors is therefore obtained.

Obtaining partial solutions from 2-view constraints First, from the estimated minors of the 2-view constraint, we can easily obtain the partial Euclidean solution based on Shapiro et al.'s reformulation of Koenderink and Van Doorn's rotation representation. Koenderink and Van Doorn's representation is probably the most appropriate parameterisation, since it distinguishes clearly between the entities which can be obtained from two views and those that can not.

Assume that the affine epipolar geometry of two views is estimated as $(a, b, c, d)(u, v, u', v')^T = 0$. From 3 ratios $a : b : c : d$, exactly 3 Euclidean parameters can be extracted. Using a scaled rotation matrix instead of $\mathbf{M}$, the following relation holds: $a : b : c : d = sr_{32} : -sr_{31} : r_{23} : -r_{13}$.

Therefore, the relative scale factor between the two views is immediately given as $s = \frac{a^2+b^2}{c^2+d^2}$.

According to Koenderink and Van Doorn, the entire rotation can be decomposed into a rotation in the image plane (assume this rotation angle to be θ) and a rotation through an angle ρ about an axis (angled at ϕ to the positive x axis) in a frontoparallel plane. The rotation matrix in terms of Koenderink and Van Doorn's $\theta - \phi - \rho$ representation [16] can be recomposed as $\mathbf{R}_{3\times 3}$:

$$\begin{pmatrix} (1-c(\rho))c(\phi)c(\phi-\theta)+c(\rho)c(\theta) & (1-c(\rho))c(\phi)s(\phi-\theta)-c(\rho)s(\theta) & s(\phi)s(\rho) \\ (1-c(\rho))s(\phi)c(\phi-\theta)+c(\rho)s(\theta) & (1-c(\rho))s(\phi)s(\phi-\theta)+c(\rho)c(\theta) & -c(\phi)s(\rho) \\ -s(\rho)s(\phi-\theta) & s(\rho)c(\phi-\theta) & c(\rho) \end{pmatrix}.$$

Therefore,

$$a : b : c : d =$$
$$s\sin\rho\cos(\phi-\theta) : s\sin\rho\sin(\phi-\theta) : -\cos\phi\sin\rho : -\sin\phi\sin\rho$$

hence, the rotation angle in the image plane is easily determined by $\tan\phi = \frac{d}{c}$, and the rotation axis modulo π out of the image plane by $\tan(\phi-\theta) = \frac{b}{a}$.

Obtaining the full solution with the 3-view constraint Up to this point, the only unknown is the rotation angle out of the image plane ρ, which is the only component that generates depth information. The one-parameter family of solutions for the rotation matrix between the two views is

$$\mathbf{R}_{3\times3}(\rho) = \begin{pmatrix} \mathbf{D}_{2\times2} + \cos\rho\mathbf{E}_{2\times2} & \sin\rho\mathbf{F}_{2\times1} \\ \sin\rho\mathbf{G}_{1\times2} & \cos\rho \end{pmatrix},$$

where $\mathbf{D}_{2\times2}, \mathbf{E}_{2\times2}, \mathbf{F}_{2\times1}$ and $\mathbf{G}_{1\times2}$ are the known quantities.

Similarly, with the second 2-view constraint, we get another one-parameter family of solutions for the rotation matrix $\mathbf{G}_{3\times3}(\rho')$ of the other 2 views in terms of the unknown rotation angle out of the image plane ρ'.

Now, it is time to use the 3-view constraint to fully determine the motion/structure. It can be easily verified that

$$t_4 : t_8 : t_{11} : t_9 = ss'(r_{11}g_{13} - g_{11}r_{13}) : ss'(r_{12}g_{13} - g_{12}r_{13}) : -s'g_{13} : sr_{13}.$$

Substituting the ratio t_9/t_{11} into t_4/t_{11} and t_8/t_{11}, we get exactly 2 linear equations in $\cos\rho$ and $\cos\rho'$,

$$\begin{cases} \frac{t_4}{t_{11}} = a\,\cos\rho + b\,\cos\rho' + c \\ \frac{t_8}{t_{11}} = a'\cos\rho + b'\cos\rho' + c' \end{cases} \text{ i.e. } \mathbf{A}_{2\times2}\begin{pmatrix}\cos\rho \\ \cos\rho'\end{pmatrix} = \mathbf{B}_{2\times1}.$$

Remarks

- This new formulation contains only $9 = 10 - 1$ independent parameters, compared with the 8 Euclidean motion parameters of the set of 3 views, it is a minimal linear parameterisation.

- The advantage of solving $\mathbf{t}$ as a whole is that it provides more information, for instance the knowledge of the ratio $t_9 : t_{11}$ which could not be recovered from 2-view constraints is the key for a linear solution in $\cos\rho$ and $\cos\rho'$. Huang and Lee had to re-compute this ratio at the very begining of the second step with the rotation composition constraint.

- Solving directly and linearly for $\cos\rho$ and $\cos\rho'$ is of great significance. On the one hand, the intrinsic two-way ambiguity is nicely expressed by the fact that

  $$\cos(-\rho) = \cos\rho.$$

  This parameterisation indeed makes a linear solution possible since the two equations are solved together. On the other hand, the only failure mode of the entire linear algorithm is the possibility that $\cos\rho > 1$ which may happen due to numerical error when $\cos\rho$ is close to 1. Since

  $$\cos\rho \approx 1 \Rightarrow \rho \approx 0,$$

  this means that actually there is almost no rotation out of the image plane. As rotation out of the image plane is the only component which contains depth information, this means that the 2D images we used do not contain the desired 3D structure, or equivalently we can report that $\rho = 0$ for the motion. Essentially, this algorithm does not have the failure mode that the factorisation method suffers in its linear version.

4.1 Comparison with related work

We first make a comparison with the existing linear algorithms of Huang and Lee and the factorisation method of Tomasi and Kanade.

Basically, there are two steps in Huang and Lee's method [7], the first step computes the coefficients of the three 2-view constraints. Any 3-view constraint was totally absent during the batch solution step and was introduced in the second step by the composition rule of rotation matrices $\mathbf{R}_{13} = \mathbf{R}_{23}\mathbf{R}_{12}$. In our new linear algorithm, the 3-view constraint has been already integrated in the first numerical step. No any other constraint was used afterwards. Since Euclidean depth information is only contained in 3-view constraints, therefore, it is dangerous to not use any 3-view constraint during the numerical step. Although 3 2-view constraints have also $9 = 3 \times 3$ independent parameters if each one is estimated individually, but they are essentially a set of 12 homogeneous parameters which breaks up into 3 sets of 4 homogeneous parameters. If we examine carefully the second step of Huang and Lee's method, it has to recompute this set of 12 homogeneous parameters (via the ratios between these 3 sets of 4 homogeneous parameters) using the 3-view constraint. In conclusion, Huang and Lee's method does not use the appropriate constraints. It will inevitably lead to the inconsistencies of the rotation matrix.

Compared with the factorization method of Tomasi and Kanade [18, 12, 22]—most suitable for redundant views, the major problem is to impose the 'metric constraint'. The linearly estimated matrix which is the product of an affine transformation and its transpose may not be positive definite, the whole Euclidean reconstruction process fails. The exact Cholesky parameterisation of the matrix introduced in [14] needs to solve simultaneous quadratic equations.

Other important work include Koenderink and Van Doorn [8]. The method consists of three steps. The first step shows that the scale change between 2 views, the rotation in the image plane around the viewing direction and the projection of the rotation axis out of the image plane can be ob-

tained with 2 views of 4 points. The second step is to parameterise the remaining Euclidean structure with the angle of the rotation out of the image plane and the 2 depths of the reference triangle, then eliminate the unknown angle to get a quadratic equation on the 2 Euclidean depths. Finally, with the third view, a second quadratic equation is obtained. Intersecting these two quadratics gives 4 possible solutions for the two depths. These intersections represent either one or two pairs of solutions that are related through a reflection in the fronto-parallel plane.

We can see that the first step of Koenderink and Van Doorn's method is similar to that of Huang and Lee [7] and Lee and Huang [9], it uses 2-view constraint to get partial solutions although Koenderink and Van Doorn's is more geometrically oriented. One major difference is that Koenderink and Van Doorn do not use the third 2-view constraint as Huang and Lee did. Unfortunately, Koenderink's method needs to intersect two quadratics for each pair of points, and can not handle all available points.

Shapiro et al. [16] extended Koenderink and Van Doorn's first step by nicely relating the Koenderink and Van Doorn's rotation representation to the affine epipolar geometry. Unfortunately they failed to get a closed form solution and adopted a non linear numerical optimisation approach for the 3-view case.

Ullman and Basri [21] and Poggio [13] considered the 3-view constraint for linear combination for recognition. Although they essentially show the equivalence with motion/structure, they do not concentrate on motion/structure recovery, in fact there is no closed form solution which allows Euclidean structure extracted directly from linear combination coefficients.

5 Experimental results

The linear method for motion/structure from 3 calibrated affine images developed in this paper has been implemented and applied to real image sequences.

We first acquired a sequence of images of a calibration pattern with a standard camera mounted on a robot. It is important to stress that the imaging conditions were not chosen to be close to affine. The triplet of images we used is shown in Figure 1. The 69 points have been automatially identified and tracked for the triplet.

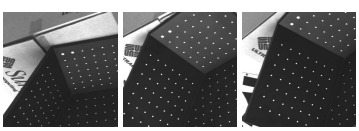

Figure 1. The triplet of images of the calibration pattern.

The 8 Euclidean motion parameters are estimated by the linear method as $s = 1.06$, $\phi = -1.51$, $\theta = -0.33$, $\rho = 0.33$ and $s' = 1.11$, $\phi' = -1.53$, $\theta' = -0.47$, $\rho' = 0.43$. The resulting shape reconstruction from these motion parameters is shown in Figure 2.

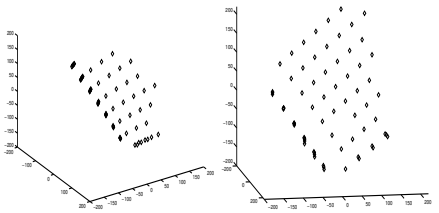

Figure 2. Two views of the resulting 3D reconstruction.

To evaluate the reconstruction quality, we did the same 3D reconstruction using a full perspective camera model, for instance the method described in [3] is used. The two reconstructions differ by a 3D similarity transformation which can be easily estimated. The normalised relative error of the Euclidean reconstruction with affine camera with respect to perspective camera one is 3.6 percent.

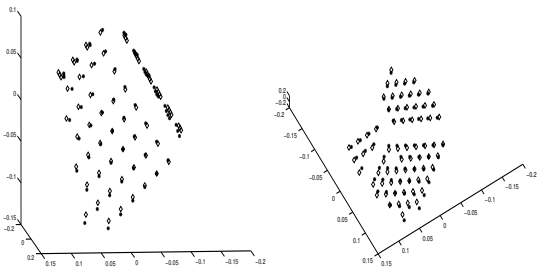

Figure 3. Two different views of the two superimposed 3D reconstructions, one uses weak-perspective camera model (marked by a square) and the other full perspective (marked by a circle).

We also tried our method on the popular hotel sequence kindly provided by Poelman and Kanade at CMU. In this sequence, the camera motion included substantial translation away from the camera and across the field of view. 197 points throughout the sequence of 181 images are automatically identified and traced. For a more detailed description of this set-up, consult [12]. The triplet of images we used are displayed in Figure 4. The resulting 3D reconstruction is shown in Figure 5.

6 Conclusion

We have introduced a unified approach to 2-view and 3-view geometric constraints of affine images, and a new

Figure 4. The triplet of hotel image sequence.

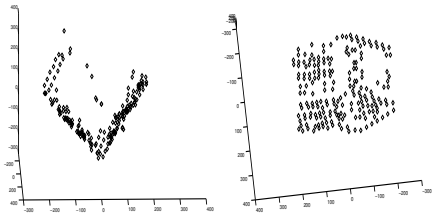

Figure 5. Two views of the 3D reconstruction of the hotel.

linear method for Euclidean motion/structure from 3 calibrated affine images. The method has been validated on real image sequences. The application of the method to novel view synthesis is underway.

What is also important for the framework of multiple affine views developed in this paper is that line segment features could also be incorporated. Line segments over 3 affine views are essentially characterised by a $2 \times 2 \times 2$ trilinear tensor [15]. It is easy to prove [24] that the 3-view constraints on the points share the 8 components of the tensor, this opens the way to combine points with line segments. The more detailed developement could be found in [24]. We can also mention that the geometric constraints of 4 affine views involving only one image coordinate can also be derived using the same framework.

Acknowledgment This work was in part supported by European CUMULI project and University of Tsukuba. We would also like to thank S. Christy for providing comparative reconstruction results.

References

[1] J.Y. Aloimonos. Perspective approximations. IVC, 8(3):179–192, August 1990.

[2] B.M. Bennet, D.D. Hoffman, J.E. Nicola, and C. Prakash. Structure from two orthographic views of rigid motion. JOSA, 6(7):1052–1069, 1989.

[3] S. Christy and R. Horaud. Euclidean shape and motion from multiple perspective views by affine iterations. PAMI, 18(11):1098–1104, November 1996.

[4] O. Faugeras. *Three-Dimensional Computer Vision - A Geometric Viewpoint*. Artificial intelligence. The MIT Press, Cambridge, MA, USA, Cambridge, MA, 1993.

[5] O. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences between n images. In ICCV95, pages 951–956, June 1995.

[6] R.I. Hartley. Lines and points in three views and the trifocal tensor. IJCV, 22(2):125–140, 1997.

[7] T.S. Huang and C.H. Lee. Motion and structure from orthographic projections. PAMI, 11(5):536–540, 1989.

[8] J. Koenderink and A. van Doorn. Affine structure from motion. JOSA, 8(2):377–385, 1991.

[9] C.H. Lee and T. Huang. Finding point correspondences and determining motion of a rigid object from two weak perspective views. CVGIP, 52:309–327, 1990.

[10] Y. Mukaigawa, Y. Nakamura, and Y. Ohta. Synthesis of arbitrarily oriented face views from two images. In ACCV95, pages 718–722, 1995.

[11] J.L. Mundy and A. Zisserman, editors. *Geometric Invariance in Computer Vision*. The MIT Press, Cambridge, MA, USA, 1992.

[12] C.J. Poelman and T. Kanade. A paraperspective factorization method for shape and motion recovery. In ECCV94, pages 97–108. Springer-Verlag, May 1994.

[13] T. Poggio and S. Edelman. A network that learns to recognize three-dimensional objects. *Nature*, 343(6255):263–266, January 1990.

[14] L. Quan. Self-calibration of an affine camera from multiple views. IJCV, 19(1):93–105, May 1996.

[15] L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. PAMI, 19(8):834–845, August 1997.

[16] L.S. Shapiro, A. Zisserman, and M. Brady. 3D motion recovery via affine epipolar geometry. IJCV, 16(2):147–182, 1995.

[17] A. Shashua. Algebraic functions for recognition. PAMI, 17(8):779–789, August 1995.

[18] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. IJCV, 9(2):137–154, November 1992.

[19] B. Triggs. Matching constraints and the joint image. In E. Grimson, editor, ICCV95, pages 338–343. June 1995.

[20] S. Ullman. *The Interpretation of Visual Motion*. The MIT Press, Cambridge, MA, USA, 1979.

[21] S. Ullman and R. Basri. Recognition by linear combinations of models. PAMI, 13(10):992–1006, 1991.

[22] D. Weinshall and C. Tomasi. Linear and incremental acquisition of invariant shape models from image sequences. In ICCV93, pages 675–682. May 1993.

[23] Z. Zhang, K. Isono, and S. Akamatsu. Euclidean structure from uncalibrated images using fuzzy domain knowledge: Application to facial images synthesis. In ICCV98, 1998.

[24] L. Quan and Y. Ohta. Notes on a general framework for 3-view based image synthesis with affine cameras. T.R. Oct. 1997.