SELF-CALIBRATION AND METRIC RECONSTRUCTION FROM SINGLE IMAGES

Ruisheng Wang, Frank P. Ferrie

Centre for Intelligent Machines, McGill University, 3480 University Street, Montreal, Quebec, Canada, H3A 2A7 - (ruisheng, ferrie)@cim.mcgill.ca

Commission V, WG V/4

KEY WORDS: Calibration, Model based, Vanishing Points, Metric Reconstruction

ABSTRACT:

It is well known that vanishing point based analysis requires images with strong perspective effects due to wide angle lenses, close objects, and often oblique viewing angles. Aerial imagery, however, normally present weak perspective effects because of long-range shootings. In this paper, we present a model-based method for reconstructing rectilinear buildings from single images. The recovery algorithm is formulated in terms of two objective functions which are based on the equivalence between the vector normal to the interpretation plane in the image space and the vector normal to the rotated interpretation plane in the object space. These objective functions are minimized with respect to the camera pose, the building dimensions, locations and orientations to obtain estimates for the structure of the scene. The comparison with the vanishing points based method indicates that our method is significantly superior over the vanishing points based method. The effectiveness of this approach is also demonstrated quantitatively through simulations and actual images.

1. INTRODUCTION

1.1 Background

3D object reconstruction from images is a common problem in computer vision and photogrammetry. In the case of only one image available, 3D reconstruction from single images has to be performed. The existing single view reconstruction methods may be roughly divided into two broad categories: geometry constraints based and model based. Here we do not include non-metric single view reconstruction methods (e.g. Hoiem et al., 2005) in these two categories. The geometry constraint based methods (e.g. Leibowitz et al., 1999) use the geometry inherent in the images (i.e. vanishing points) to derive the camera calibration information and result in a 3D reconstruction; while most model-based methods (e.g. Debevec et al, 1996) recover model parameters through a model-to-image fitting algorithm which involves a minimization of total disparity errors between observed edges and projections of the reconstructed lines.

Vanishing points are defined as points at which the extensions of parallel lines appear to converge in the perspective view of the image. The limitations of vanishing points based methods are obvious. As vanishing points are points in the image at infinity, slight inaccuracy in the measurements of lines will result in large errors in the positions of calculated vanishing points. Automatic methods improve the accuracy of vanishing detection but often require sufficient straight lines which are detectable in the images. In general, the images may, however, contain very few straight lines. Further, no existing edge detection algorithms can provide only useful edges reliably from images of a common scene; human intervention is always needed in those automatic methods. From a practical point of view, manual digitization of straight lines in the images is often involved. In this situation, the vanishing points based methods do not work well, in particular for those images with weak perspective effects (e.g. aerial images).

In this paper, we present a model-based method to reconstruct rectilinear buildings up to a scale factor from single images. The difference from previous model-based methods is that our method does not require a model-to-image fitting algorithm, and therefore avoid a minimization procedure. The method is based on manual feature correspondence between pre-defined parameterized 3D model edges and corresponding image edges. The algorithm then automatically recovers camera pose and model dimensions.

To our knowledge, the comparison between vanishing points based methods and model-based methods are rarely reported. In this paper, we also compared the performance of the vanishing points based method (Zhang et al. 2001) with our model-based method using identical synthetic and real data. The quantitative analysis results indicate that our method is significantly superior over the vanishing points based method in terms of feasibility for various images with strong or weak perspective effects.

1.2 Related Work

1.2.1 Vanishing Points Based Methods: The existing vanishing points detection methods may include manual detection, using Hough Transform (e.g. Tuytelaars 1997), searching over Gaussian sphere (e.g. van den Heuvel, 1998), and using projective geometry (e.g. Birchfield, 1998). Most of automatic vanishing points detection methods are not only computational intensive but also require human interaction, which are hard to reach operational level. Manual detection of vanishing points satisfies operational level but suffer problem that the determined vanishing points may not be accurate. There also exist a bunch of papers on vanishing points based 3D reconstruction (e.g. Guillou et al., 2000). Since our method is not in this route, we do not include a detailed review of the

methods here. Interest readers may refer the paper (Guillou et al., 2000) for more information.

1.2.2 Model-Based Methods: We regard model-based methods (e.g. Debevec et al. 1996, Wang and Ferrie, 2008) as evolution from line-based methods (e.g. Liu et al. 1990, Kumar and Hanson 1994). The model-based methods use structure information inherent in the objects which is ignored in the line-based approaches. Early attempt (Liu et al. 1990) solved for the camera rotation first and then the camera translation using both lines and points correspondence. They considered three camera rotation angles as obtained from a nominal orientation by small perturbations, e.g. 0 degrees. Based on this assumption, their algorithm only works if the three camera Euler rotation angles are less than 30 degrees. Kumar and Hanson (1994) solved for the rotation and translation simultaneously by adapting an iterative technique formulated by Horn (1990). They also reported that the initial rotation estimates for some data sets must be within 40 degrees for all the three Euler angles representing the rotation. When initial estimates for rotation and translation are not available, they sampled rotation space, and each of the samples was used as an initial estimate for the rotation estimation by a method akin to Liu et al. (1990). The estimated rotation and translation based on the rotation samples are then used as initial estimates for solving the camera rotation and translation simultaneously. Taylor and Kriegman (1995) estimated both the camera positions and the structure of the scene from multiple images. Based on a random initial estimate of rotation, the translation and model parameters are computed as initial inputs for the subsequent model-to-image fitting procedure. If the disparity between predicted edges and the observed edges is smaller than some preset threshold, the minimum is accepted as a feasible estimate. Debevec et al. (1996) argued that if the algorithm begins at a random location in the parameter space, it stands little chance of converging to correct solution. They developed a method to directly compute a good initial estimate for the camera positions and model parameters, and then use those estimates as initial inputs for the subsequent model-to-image fitting process.

Our approach builds on this line of work. Described is a two-step iterative scheme for recovering camera orientation that, unlike existing methods, does not require a good initial guess for the rotation. Instead, the good initial estimate for the rotation is computed directly by using coplanarity constraints. The camera translation and predefined model parameters are determined based on the calculated rotation through a linear least squares minimization. The 3D reconstruction of buildings is based on the recovered camera pose and the assumption of flat terrain. Unlike existing methods, our method does not require a model-to-image projection process, and is particularly suitable for oblique images with large shooting angles in urban environments.

2. THE METHOD

2.1 Notation

Figure 1 shows how a straight line segment, model edge 67, in a cube model (building 1) projects onto the image plane of a camera. The coordinates of two endpoints of the projected image edge 67 in the camera coordinate system can be represented as $\{(x_1, y_1, -f), (x_2, y_2, -f)\}$. The camera position relative to the object coordinate system is represented in terms of a rotation matrix R and a translation vector t. The straight line 67 can be defined by a pair of vectors (v, u) in the object coordinate system where v represents the direction of the line and u represents a point on the line. m is normal vector of the projection plane defined by the two lines (C_6, C_7) and camera centre C in the camera coordinate system. The coplanar constraints derived in (Taylor and Kriegman, 1995) are outlined in the following. The fundamental relation of the imaging geometry can be represented by the equation (1),

$$m = R(v \times (u - t)) \tag{1}$$

Equation (1) is based on the fact that the 3D model lines (e.g. line *67*) in the camera coordinate system must lie on the projection plane formed by lines (*C6*, *C7*) and camera centre *C*.

$$m^T R v = 0 \tag{2}$$

$$m^T R(u - t) = 0 \tag{3}$$

Equations (2) and (3) are deduced from equation (1), which shows that the determination of camera rotation R can be independent from the estimation of camera position t and model parameters. Note v becomes a known vector in the object-centered coordinate system which is parallel to the Y axis, while u can be represented by the model parameters.

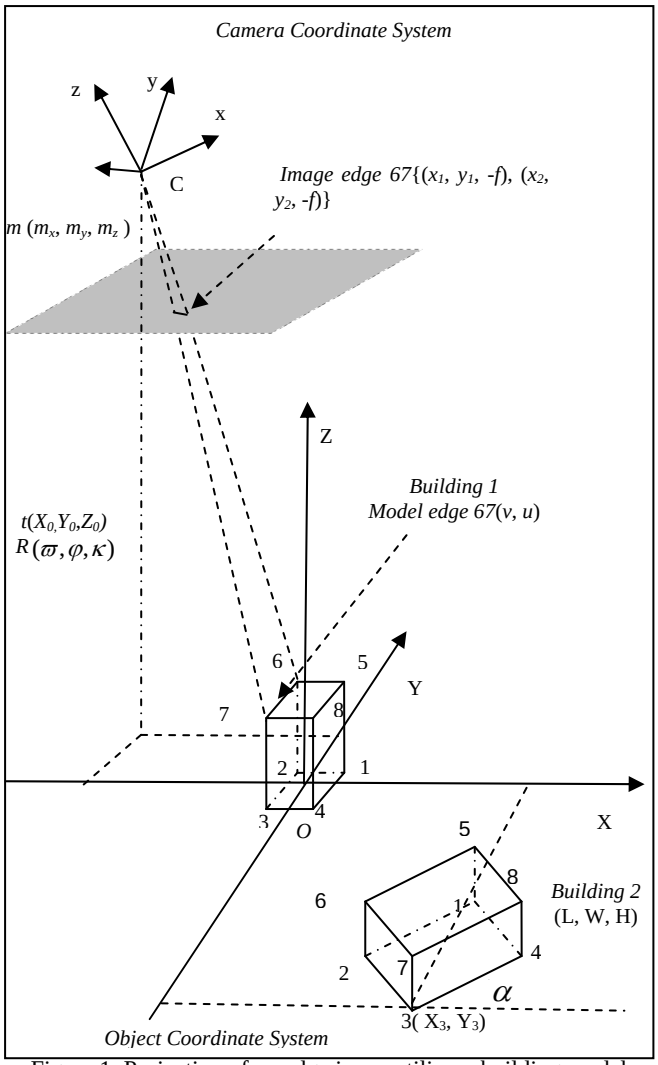

Figure 1. Projection of an edge in a rectilinear building model onto a camera's image plane and spatial relationship of buildings

The normal vector m can be defined by the intersection of the projection plane C_{67} with the image plane as shown in Figure 1 and represented in the equation (4),

$$m_x x + m_y y - m_z f = 0 \tag{4}$$

where m_x, m_y, m_z are the coordinates of the normal vector in the camera coordinate system and f is the focal length of the camera; x and y are points on the image edge. Given an observed image of edge 67, the observed normal vector m' can be obtained by the equation (5),

$$m' = (x_1, y_1, -f)^T \times (x_2, y_2, -f)^T \tag{5}$$

The location and orientation of the building 2 can be represented by a building vertex (e.g. vertex 3 (X_3, Y_3) in the Figure 1), a building orientation along the X axis (e.g. the α in Figure 1), and the building's dimension of length, width, and height (e.g. L, W, H in Figure 1). Those unknown parameters are solved in the metric reconstruction stage.

2.2 Recovery Algorithm

The recovery algorithm takes as input, a set of correspondences between edges in the models and edges in the image. The correspondences are performed manually. The algorithm then automatically recovers camera pose and model dimensions, consisting of self-calibration and metric reconstruction. In the first step, the focal length is firstly obtained from image EXIF tags. The camera pose and the model parameters are recovered with respect to an object-centred coordinate system. In the second step, the spatial relationship of buildings is represented by three intrinsic parameters (building length, width, and height) and three extrinsic parameters (a building vertex location and building orientation). Those parameters can be determined by using model-to-image correspondence and the recovered camera pose.

2.2.1 Self-Calibration

The self-calibration requires more than three line correspondences between the pre-defined model edges and the image edges, which consists of initial estimate of camera rotation, refinement of camera rotation, and determination of camera translation and model dimensions.

Initial Estimate of Camera Rotation

The objective function of obtaining initial estimates for camera rotation is formulated according to the Equation (2) as shown in the Equation (6),

$$O_1 = \sum_i^n (m_i^T R v_i)^2 \tag{6}$$

where i is the number of the model edges, n is the total number of the employed model edges, m_i and v_i are the corresponding normal vector and direction of the model edge, R is 3x3 camera rotation matrix. By summing up the extents to which the rotation R violates the constraints arising from Equation (2), the objective function can be minimized to obtain initial values for the camera rotation

Refinement of Camera Rotation

Once initial camera rotation is obtained, a non-linear technique based on Gauss-Newton method is applied to the minimization problem. The direct calculation of Jacobian matrix of the objective function O_1 is complex. To simplify the linearization of O_1, we rewrite the rotation matrix R as a multiplication of three sequential rotations, and compute the first derivative for each rotation angle. The Jacobian matrix of O_1 can then be formed as,

$$J_{\substack{\varpi\varphi\kappa \\ n\times 3}} = \begin{bmatrix} m_1^T R_\varpi' v_1 & m_1^T R_\varphi' v_1 & m_1^T R_\kappa' v_1 \\ \vdots & \vdots & \vdots \\ m_n^T R_\varpi' v_n & m_n^T R_\varphi' v_n & m_n^T R_\kappa' v_n \end{bmatrix},$$

where

$$R_\varpi' = R\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},\ R_\varphi' = \begin{bmatrix} 0 & 0 & -\cos\kappa \\ 0 & 0 & \sin\kappa \\ \cos\kappa & -\sin\kappa & 0 \end{bmatrix} R,$$

$$R_\kappa' = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} R$$

Given the three initial camera rotations obtained from the previous step, the Gauss-Newton algorithm computes accurate estimates of the camera rotations within 2-3 iterations.

Determination of Camera Translation and Model Dimensions

The objective function for determining camera translation and model dimensions is formulated according to Equation (3) as shown in Equation (7),

$$O_2 = \sum_i^n (m_i^T R(u_i - t))^2 \tag{7}$$

where i is the number of the model edges, n is the total number of the employed model edges, m_i, and u_i are the corresponding normal vector and point on the model edge. In the case of rectilinear buildings, the minimization of the objective function O_2 is a constrained quadratic form minimization problem, and can be solved through a set of linear equations. It is also important to keep in mind that the resulting dimensions of the scene and camera translations are up to a scale factor.

2.2.2 Metric Reconstruction

The metric-reconstruction also requires more than three line correspondences between the pre-defined model edges and the image edges, which consists of initial estimate of building orientation, refinement of building orientation, and determination of building dimensions and location.

Initial Estimate of Building Orientation

The three directions of model edges, v_1 (e.g. model edge 67 of the building 2 in Figure *1*), v_2 (e.g. model edge 78), and v_3 (e.g. model edge 27), can be represented as shown in Equation (8).

$$v_1 = \begin{pmatrix} -W\sin\alpha \\ W\cos\alpha \\ 0 \end{pmatrix},\ v_2 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} v_1,\ v_3 = \begin{pmatrix} 0 \\ 0 \\ H \end{pmatrix} \tag{8}$$

Since the value of building dimensions (*W*, *H*) does not affect the building orientation, the only unknown parameter of building orientation is α . At this stage, the camera orientation is known. The objective function of Equation (6) is employed to obtain an initial estimate for α . The minimization of the objective function (6) sums the extents to which the model orientation *v* violates the constraints arising from Equation (2).

Refinement of Building Orientation

Once initial building orientation is obtained, a non-linear technique based on the Gauss-Newton method is applied to the minimization problem. Based on Equation (9), the minimization is straightforward since there is only one unknown parameter α ,

$$-a_{i1}^T W \sin\alpha + a_{i2}^T W \cos\alpha = 0 \quad (9)$$

where $a_i^T = m_i^T R$

Given the initial building orientation obtained from the previous step, the Gauss-Newton algorithm computes the accurate building orientation within 2-3 iterations.

Determination of Building Dimensions and Location

The building dimensions and location is determined by minimizing objective function O_2 as shown in Equation (7). In this stage, image norm vectors *m*, camera rotation *R* and translation *t* are known. The unknowns are points u_i on the model edges which are expressed as linear functions of building dimension and location, and can be solved through a set of linear equations. The same way can be employed to reconstruct more buildings.

2.3 Vanishing Points Based Algorithm

The method (Zhang et al., 2001) proposed an adjustment model for computing vanishing points, and then determined camera focal length as well as camera orientation using calculated vanishing points. Given a dimension of the cubic building, the camera translation and other two dimensions of the building were determined. We also extended this approach to deal with multiple buildings using the topological representation as shown in Figure 1. Then we evaluated effectiveness of our method with this extended vanishing points based method. Due to the space limit, we do not include the details of their methods in this paper.

3. EXPERIMENTAL RESULTS

This section describes a series of experiments that were carried out in order to evaluate the effectiveness of proposed algorithm and the vanishing points based algorithm. Simulation is first used to systematically vary key parameters such as the camera parameters and measurement of the image segments; thereby enabling us to characterize the degradation of the algorithms in extreme situations. Real examples are shown to gauge practical results.

3.1 Simulation Experiments

The synthetic image data was generated with a virtual camera and two 3D cubic building models as shown in Figure 1. The camera parameters are listed in the Table 1, assuming that the image centre lies at the centre of the image frame. Table 2 shows information about the two building dimensions, locations, and orientations.

Focal Length (*m*)	0.0798	Pixel Size (*um*)	12
X_0(m)	-500.672	ω (∘)	50.346
Y_0(m)	100.317	ψ (∘)	3.582
Z_0(m)	650.783	κ (∘)	2.787

Table 1. Camera intrinsic and extrinsic parameters

	Building 1	Building 2	
Length (m)	40	26.41	
Width (m)	20	20.92	
Height (m)	30	22.31	
Orientation along X axis (∘)	0	α = 30.856	
Location of a building model vertex (m)		X_3	100.512
		Y_3	-200.217

Table 2. Building parameters of dimensions, locations and orientations

We evaluated how errors in measurements of image segments as well as camera parameters influence accuracy of the recovered camera pose and building model dimensions for both methods. In the following tables, entries in rows with "V" correspond to experiments from the vanishing points based method, with "M" correspond to experiments from our model based method. Entries in column with "0" correspond to experiments with correct image measurements or camera parameters.

Errors from image noise

A uniformly distributed random image error is added to the endpoints of the image segments. Entries in columns with 0 random errors correspond to the experiments with the image segments without errors. Table 3 shows that the reconstruction errors increase as the random image errors are increased for both methods. However, the vanishing points based method is extremely unstable to those random errors in the endpoints of the image segments. The results from vanishing points based method shows a very bad reconstruction. The reason is small errors in endpoints of the image segments cause huge errors in the determination of vanishing points. Thereby, large errors in vanishing points cause huge errors in the resulted camera pose and buildings. While the model based approach is relatively stable to those random errors. The errors in the endpoints have much less effect on the accuracy of the reconstruction compared with vanishing points based method.

		Noises in endpoints of image segments (pixel)			
Camera Pose		0	5	10	15
ω (°)	V	50.346	48.183	40.773	30.269
	M	50.346	50.185	50.675	49.630
ψ (°)	V	3.582	4.327	8.317	13.104
	M	3.582	3.858	4.303	1.336
κ (°)	V	2.787	0.782	24.080	32.333
	M	2.787	2.281	2.074	6.885
X_0(m)	V	-500.672	-485.999	-557.189	-580.263
	M	-500.672	-508.195	-488.357	-527.336
Y_0(m)	V	100.317	135.713	-14.467	-121.659
	M	100.317	97.641	95.361	92.791
Z_0(m)	V	650.783	643.552	453.882	451.424
	M	650.783	673.415	658.684	601.905
Building 1					
W(m)	V	20.000	22.023	16.302	14.941
	M	20.000	20.672	20.147	18.995
H(m)	V	30.000	31.710	21.055	21.847
	M	30.000	30.829	32.836	28.907
Building 2					
α (°)	V	30.856	47.744	68.578	66.743
	M	30.856	31.214	31.011	29.335
X_3(m)	V	100.512	293.388	496.548	182.584
	M	100.512	101.313	98.044	100.513
Y_3(m)	V	-200.217	-297.293	-128.671	-111.903
	M	-200.217	-201.372	-195.254	-191.698
L(m)	V	26.413	33.982	54.871	29.901
	M	26.413	26.433	25.961	25.260
W(m)	V	20.927	46.757	64.184	21.700
	M	20.927	21.059	20.650	20.089
H(m)	V	22.315	24.475	27.667	15.535
	M	22.315	22.4.7	21.821	20.245

Table 3. Comparison of the vanishing points based method with the model based method using the same noises in endpoints

Errors from image centre offsets

The influence of the incorrect image centre on both methods was analyzed introducing an error of 5, 10, 15 pixels on *x* and *y* coordinates of the image centre. In Table 4, entries in column with 0 pixel offsets correspond to the experiments with correct image centre. The experimental results show that both methods are insensitive to errors in the image centre offsets, which validate the feasibility of the approximation that the principle points lay at the image centre. However, without considering the noise in the endpoints of image segments, the accuracy of the vanishing points based method is slightly better than model based one with the same amount of image centre errors. This can be partially explained by the fact that the rotation and translation constraints are weak constraints when used separately. Small errors in the rotation are amplified into large errors in the translation, and subsequently affect the resulted building parameters.

3.2 Real Data Experiments

We take pictures using a Canon PowerShot SD750 digital camera. Figure 2a (3072X2304 pixels) is two boxes, and Figure 2b (3072X2304 pixels) is the Burnside Hall at the downtown campus of McGill University, Montreal. The measured image edge features are those black lines digitized using mouse. Table 5 shows that accuracy of 3D reconstruction from Figure2a using the model-based method is much higher than those from the vanishing points based methods, especially in the dimension of the height.

		Errors in the image centre (pixel)			
Camera Pose		0	5	10	15
ω (°)	V	50.346	50.324	50.302	50.280
	M	50.346	50.296	50.247	50.198
ψ (°)	V	3.582	3.582	3.582	3.582
	M	3.582	3.538	3.495	3.452
κ (°)	V	2.787	2.790	2.793	2.795
	M	2.787	2.865	2.942	3.018
X_0(m)	V	-500.672	-500.934	-501.197	-501.459
	M	-500.672	-501.341	-502.002	-502.654
Y_0(m)	V	100.317	100.392	100.468	100.543
	M	100.317	100.404	100.490	100.574
Z_0(m)	V	650.783	650.579	650.376	650.172
	M	650.783	649.026	647.276	645.532
Building 1					
W(m)	V	20.000	20.009	20.018	20.027
	M	20.000	19.989	19.978	19.967
H(m)	V	30.000	29.973	29.945	29.918
	M	30.000	29.879	29.759	29.640
Building 2					
α (°)	V	30.856	30.821	30.786	30.751
	M	30.856	30.819	30.782	30.745
X_3(m)	V	100.512	100.403	100.294	100.185
	M	100.512	100.529	100.546	100.564
Y_3(m)	V	-200.217	-200.298	-200.379	-200.461
	M	-200.217	-200.156	-200.096	-200.035
L(m)	V	26.413	26.422	26.431	26.441
	M	26.413	26.419	26.424	26.430
W(m)	V	20.927	20.928	20.929	20.930
	M	20.927	20.921	20.916	20.910
H(m)	V	22.315	22.296	22.278	22.258
	M	22.315	22.242	22.171	22.099

Table 4. Comparison of the vanishing points based method with the model based method using the same image centre errors

(a) (b)

Figure 2. (a) Two boxes with digitized edges superimposed (b) Burnside Hall with digitized edges superimposed

The building in Figure2b is an irregular cube but we use a rectilinear building model to approximate it, which induces measurement errors, especially in the corners of the building. Besides, the occlusions caused by snow make the accurate measurement of the building top and bottom difficult. Under all of this noise, however, we still achieve reasonable results using model based approach. The computed dimensions of Burnside Hall are 35.44, 34.92, and 53.33 meters respectively, as compared to the model dimensions obtained from DWG file of 35.44, 32.42, 50.00 meters. While the vanishing points based

method does not work properly in this situation. Figure 3 shows the recovered camera pose and wire frame of the Burnside Hall using MatLab.

Unit: mm			Dimensions measured using a ruler	Dimensions computed from the image
Left Box	Width	V	71.1	73.1
		M		70.9
	Height	V	123.1	97.9
		M		122.6
Right Box	Length	V	72.1	68.8
		M		71.7
	Width	V	50.6	47.6
		M		49.6
	Height	V	15.3	6.3
		M		14.8

Table 5. Comparison of the vanishing points based method with the model based method using real image Fig2.a

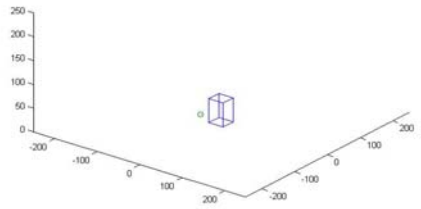

Figure 3. Visualization of the recovered camera pose and wire frame of the Burnside Hall

4. CONCLUSIONS

This paper presented a method to recover 3D rectilinear building models from single monocular images. The method uses the correspondences between predefined 3D models and their corresponding 2D images to obtain camera pose as well as parameters of 3D building models. The camera orientation is first recovered followed by solving translation and the first building model dimensions. The direct computation of the initial estimate for camera rotation effectively solved problems in the previous approaches (e.g., Taylor and Kriegman, 1995), and the determination of camera pose as well as the first building model dimensions are much simpler than the previous methods (e.g., Debevec et al., 1996). Under the assumption of flat terrain, more 3D building models can be reconstructed based on recovered camera pose through model-to-image correspondence.

Simulation experiments were carried out in order to investigate how the accuracy of the algorithm would be affected as different parameters were varied. The comparison using identical synthetic and real data shows that our method is significantly superior over the vanishing points based method. The experiments also show that our method robustly and accurately estimates camera pose and building dimensions provided that accurate image measurements are available.

REFERENCES

Camillo J. Taylor and David J. Kriegman, 1995. Structure and motion from line segments in multiple images, *IEEE Transaction on Pattern Analysis Machine Intelligence*, 17(11).

Caprile B. and Torre, V., 1990. Using vanishing points for camera calibration, *International Journal of Computer Vision* (4), No. 2, pp.127 -140

E. Guillou, D. Meneveaux, E. Maisel, and K. Bouatouch, 2000. Using vanishing points for camera calibration and coarse 3-D reconstruction from a single image, *Visual Computer*, vol. 16, no. 7, pp. 396–410.

Debevec P., C. Taylor, and J. Malik, 1996. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry and Image-based Approach, *SIGGRAPH*, pp. 11 -20.

F.A. van den Heuvel, 1998. Vanishing point detection for architectural photogrammetry. In *International Archives of Photogrammetry and Remote Sensing* Vol. XXXII part 5, pages 652–659.

Horn, B.K.P, 1990. Relative Orientation, *International Journal of Computer Vision* Vol. 4, No. 1, pp. 59—78.

Hoiem D., A. Efros, and M. Hebert, 2005. Automatic photo pop-up. In *ACM SIGGRAPH*.

Kumar, Rakesh, and Hanson, Allen R, 1994. Robust methods for estimating pose and a sensitivity analysis. *CVGIP: Image Understanding*, v 60, n 3, p 313-342.

Leibowitz D, Criminisi A, Zisserman A, 1999. Creating architectural models from images. In: Brunet R, Scopigno R (eds) *Eurographics*, Vol 18, pp 39–50, Blackwell Publishers.

Liu, Y.; Huang, T.S.; Faugeras, O.D., 1990. Determination of Camera Location from 2-D to 3-D Line and Point Correspondences, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1), Page(s):28 – 37

S. Birchfield, 1998. An introduction to projective geometry. Unpublished. http://vision.stanford.edu/˜birch/projective.

T. Tuytelaars, M. Proesmans, and L. Van Gool, 1997. The cascaded hough transform. *The proceedings of International Conference on Image Processing*.

Wang, R. and F. Ferrie, 2008. Camera Localization and Building Reconstruction from Single Monocular Images. The *IEEE Computer Vision and Pattern Recognition* (CVPR) *Workshop on Visual Localization for Mobile Platforms*, Anchorage, Alaska.

Zhang Z., Zhang J., Zhang S., 2001. 3D Building Reconstruction from Single Images. *Workshop on Automatic Engineering Surveying of Channel.*