EFFICIENT LINE MATCHING BY IMAGE SEQUENTIAL ANALYSIS FOR URBAN AREA MODELLING

Y. Kunii*, H. Chikatsu

Tokyo Denki University, Department of Civil and Environmental Engineering,
Hatoyama, Saitama, 350-0394, Japan - (kunii, chikatsu)@g.dendai.ac.jp

Commission Youth Forum

KEY WORDS: Urban, Extraction, Sequences, Matching, Modelling, Visualization

ABSTRACT:

3D city modelling from airborne imagery includes mainly two parts: (1) image processing procedures and (2) 3D modelling for man-made objects such as buildings, roads and other objects. Line extraction and stereo matching are usually utilized as an image processing procedures. However, there are some issues for automatic man-made object modelling. In particular, spatial data acquisition of buildings are important for reliable city modelling.
With this objective, this paper focuses especially on efficient line matching method using optical flow and trifocal tensor. Furthermore, line matching in general stereo matching methods were also investigated, and performance of the proposal line matching method was compared with these general methods in this paper.

1. INTRODUCTION

Recently, efficient spatial data acquisition and visualization have been receiving more attention from the view point of city planning, city regeneration, telecommunications, environmental and energy problems. Generally, in order to perform object modelling using digital image, line or feature extraction and stereo matching are performed, and many matching methods such as area based matching, future based matching have been proposed. In particular, line gives important information for building extraction, and satisfied 3D results are depend on rigorous line extraction and matching.
With this motive, the authors have been concentrating on developing an efficient line matching procedure for man-made object modelling using high vision imagery (Kunii and Chikatsu, 2003). The line matching is comprised of line extraction and line tracking, line extraction was performed by Canny operator (Canny, 1986), and line matching was performed using optical flow estimation and epipolar matching. However, more efficient line matching is needed for automatic 3D modelling due to fragment or multiple lines during the line matching procedure.
In these circumstances, this paper focuses on more efficient line matching method using trifocal tensor (Beardsley, et al., 1996). The trifocal tensor is geometric relation of 3 images, and useful for point or line matching of multiple image. Therefore, the line matching for the high vision imagery could be performed more efficiently.
Furthermore, line matching by general stereo matching methods were also investigated in this paper, and performance of the proposal line matching method was compared with these general methods.
Finally, 3D modelling for the urban area was performed by the line information in this paper.

2. HIGH VISION IMAGERY

The high vision imagery in this paper was taken by high definition television (HDTV) format. This paper reports the 3D modelling method using the high vision imagery which obtained from a helicopter at urban district of Kobe-city, Hyogo, Japan. Table 1 shows the major components of the high vision imagery, and Figure 1 shows the first frame imagery.

Table 1. Major components of the high vision imagery

Height	300m
Resolution	1920×1080 (pixel)
Number of Frames	50 frame

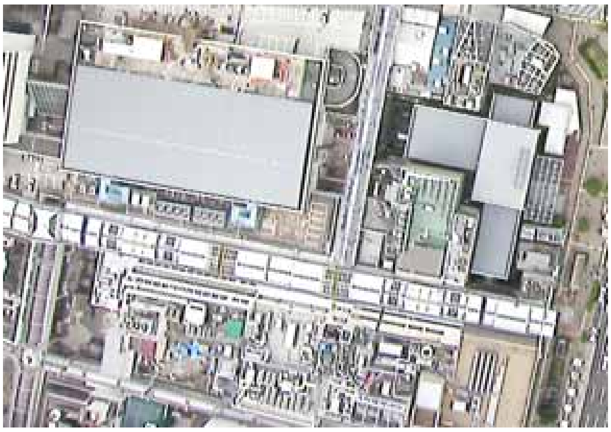


Figure 1. First frame imagery

3. LINE MATCHING BY OPTICAL FLOW

The line matching was performed by line extraction and line tracking using optical flow. Detail procedures of the line matching method are as follows.

3.1 Line Extraction

Line extraction was performed by Canny operator with 2 threshold values which called the height and reliability of edge. The height of edge is a variation of the gray level around at a interest point, and the reliability is an index for representing influence of noise. The height h and the reliability r are calculated by following equation.

$$h(x,y) = \sqrt{h_x^2(x,y) + h_y^2(x,y)} \tag{1}$$

$$r(x,y) = \frac{h(x,y)}{2\sigma_0(x,y)} \tag{2}$$

where,

h_x, h_y : variation of gray level for each direction (x,y)

x, y : image coordinate of interest point

σ_0 : variance of gray level around at interest point

These threshold values were set as h = 10 and r = 0.1 in this paper. Furthermore, both ends of these extracted edges were connected by straight lines. Figure 2 shows the extracted lines by the method for the first image (277 lines).

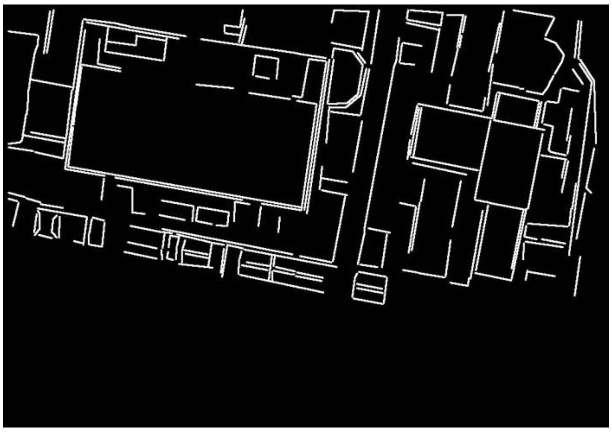


Figure 2. Line extraction

3.2 Optical Flow Estimation

In order to perform line matching, both ends for each extracted line were tracked by optical flow. Although many optical flow estimation methods have been proposed, Lucas-Kanade method (Lucas and Kanade, 1981) which is capable of correct and fast procedure was adopted in this paper. The optical flow by Lucas-Kanade method (u, v) is calculated by following equation and estimated optical flow is shown in Figure 3.

$$u = \frac{\sum_w \frac{\partial I}{\partial x} \cdot [J(p) - I(p)]}{\sum_w \left(\frac{\partial I}{\partial x}\right)^2}, v = \frac{\sum_w \frac{\partial I}{\partial y} \cdot [J(p) - I(p)]}{\sum_w \left(\frac{\partial I}{\partial y}\right)^2} \tag{3}$$

where,

$I(p) = I(x,y,t), \quad J(p) = I(x,y,t+\delta t)$

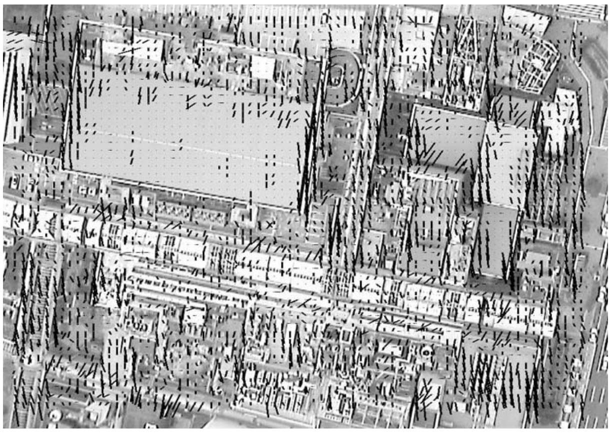


Figure 3. Optical flow estimation

4. LINE MATCHING BY TRIFOCAL TENSOR

The unmatched lines by the optical flow estimation were corrected by trifocal tensor in this paper. Details of the line matching by trifocal tensor are as follows.

Trifocal tensor is geometric relation of 3 images which contained the same objects from different perspectives. The trifocal tensor is expressed by 3 square matrixes (3×3), these 3 matrixes are $\mathbf{T_1}$, $\mathbf{T_2}$ and $\mathbf{T_3}$, components of these matrixes are t_{1ij}, t_{2ij} and t_{3ij}, and image coordinates of matched points for these 3 images are (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3). Thus, following equations are obtained by the geometric relation.

$$\begin{aligned}
-z_2z_3g_{22} + z_2y_3g_{23} + y_2z_3g_{32} - y_2y_3g_{33} &= 0 \\
z_2z_3g_{21} - z_2x_3g_{23} - y_2z_3g_{31} + y_2x_3g_{33} &= 0 \\
z_2z_3g_{12} - z_2y_3g_{13} - x_2z_3g_{32} + x_2y_3g_{33} &= 0 \\
-z_2z_3g_{11} + z_2x_3g_{13} + x_2z_3g_{31} - x_2x_3g_{33} &= 0
\end{aligned} \tag{4}$$

where,

$g_{ij} = x_1t_{1ij} + y_1t_{2ij} + z_1t_{3ij}$

These 4 equations are generated by one conjugated point of these 3 images. The trifocal tensor has 27(=3×3×3) unknown parameters which can be calculated by more than the same number of equations. Therefore, more than 7 points needed to be conjugated between these 3 images for acquisition of the trifocal tensor. Consequently, the unmatched points in the third image are calculated by the above equation.

5. RESULTS OF LINE MATCHING

In order to evaluate performance of the proposal line matching method, line matching in general stereo matching methods such as LSM (Gruen, 1985), probabilistic relaxation (Rosenfeld, et al, 1976) and area correlation (Schenk, 2001) was also investigated, and performance of the proposal method was compared with these general methods. Table 2 shows results of line matching by each method. The line matching by proposal method could be performed efficiently more than other general methods. Consequently, the optical flow estimation and the trifocal tensor is useful method for line matching by image sequences. Figure 4 shows result of line matching by the proposal method.

Table 2. Results of line matching

Method	Number of lines	Matching rate (%)
(a)	196	70.8
(b)	137	49.5
(c)	173	62.5
(d)	60	21.7

(a) Proposal method
(b) LSM method
(c) Probabilistic relaxation method
(d) Area correlation method

Figure 4. Result of line matching by proposal method

6. EPIPOLAR MATCHING

The line matching was performed efficiently by the above procedures. However, these procedures can not apply for all necessary lines due to fragment or multiple. Therefore, the unmatched lines were corrected using epipolar matching.
The epipolar matching was performed using epipolar lines for the first and last image. In order to estimate epipolar lines, relative orientation was performed by coplanarity condition using the first and last image. The both ends for the each matched lines were used as pass points, and the orientation parameters (φ_1, κ_1, ω_2, φ_2, κ_2) were determined. After the orientation, geometric correction of the first and last image was performed using the orientation parameters. Consequently, epipolar lines were estimated.
Furthermore, in order to perform stereo matching using these epipolar lines efficiently, stereo matching was performed by above 3 general methods, and performances of each method were compared. As a result, LSM method realized efficient stereo matching more than other 2 methods which shown in Table 3 and Figure 5. Consequently, LSM method was adopted for the epipolar matching in this paper.

Table 3. Results of epipolar matching

Method	Number of lines	Matching rate (%)
(a)	254	91.7
(b)	251	90.6
(c)	235	84.8

(a) LSM method
(b) Probabilistic relaxation method
(c) Area correlation method

(a) LSM

(b) Probabilistic relaxation

(c) Area correlation

Figure 5. Results of epipolar matching

7. 3D MODELLING

The line information for 3D modelling can be acquired efficiently by the method in the previous chapter. However, each rooftop of building in the urban area is needed to be recognized for 3D modelling. Therefore, rooftops recognition was performed by morphological opening procedure, and the extracted rooftops were conjugated with the matched lines in this paper (Kunii and Chikatsu, 2003). Figure 6 shows the result of the opening procedure for the first image.

Figure 6. Result of rooftops recognition

Furthermore, in order to perform 3D modelling for the urban area, camera calibration for the first image and the last image were performed by combined adjustment (Chikatsu and Kunii, 2002). Therefore, 3D data for the urban area could be calculated efficiently.
Finally, 3D modelling for the urban area was performed by following procedure: (1) side surfaces for each building were constructed using the 3D data for the both ends of the matched line, (2) the recognized rooftops were put on the side surfaces. Figure 7 shows the 3D model for the urban area.

8. CONCLUSION

This paper investigates mainly 3 issues regarding 3D modelling for urban area using image sequences: (1) efficient and robust line matching method using optical flow and trifocal tensor, (2) performance evaluation of the proposal line matching, (3)more efficient epipolar matching, and followings are main results were obtained:

+ Line matching was improved by trifocal tensor.
+ Proposal line matching method was efficiently more than other general methods.
+ Efficient epipolar matching was performed by LSM.

Thus, it is concluded that the line matching method comprised optical flow, trifocal tensor and epipolar matching is useful method for 3D modelling. However, there are still the following issues to be resolved before this method becomes operational.

+ Recognition of Complicated rooftops.
+ Texture mapping.

Acknowledgement
The high vision imagery used in this paper were taken by Aero Asahi Corporation. The authors would like to thank Aero Asahi Corporation.

References from Journals:
Canny, J., 1986. A computational approach to edge detection, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.PAMI-8, No.6, pp.679-697.

Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique, *South Africa Journal of Photogrammetry, Remote Sensing and Cartography*, Vol.14, No.3, pp.175-187.

Rosenfeld, A., Hummel, R. and Zucker, S., 1976, Scene labeling using relaxation operation, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 6, No. 6, pp.420-433.

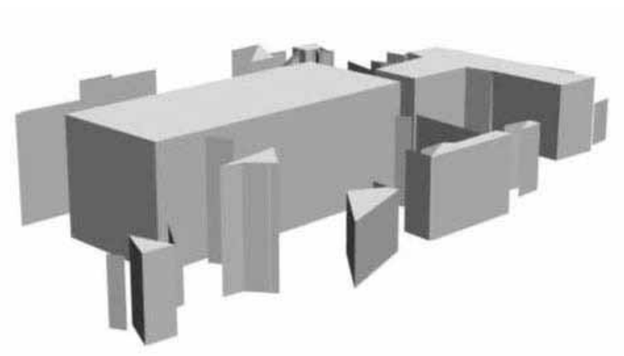
References from Books:
Schenk, T., 2001 *Digital Photogrammetry*, TerraScience, Ohio, USA.

References from Other Literature:
Beardsley, P., Torr, P. and Zisserman, A., 1996. 3D model acquisition from extended image sequences, *4th European Conference on Computer Vision*, pp.683-695, Cambridge, UK.

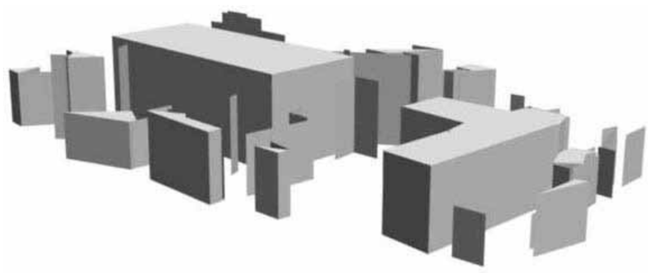
Chikatsu, H. and Kunii, Y., 2002. Performance evaluation of recent high resolution amateur cameras and application to modeling of historical structure, *Internationals Archives of Photogrammetry and Remote Sensing*, Vol.XXXIV Part5, pp.337-341, Corfu, Greece.

Kunii, Y. and Chikatsu, H., 2003, Building extraction and modeling in urban area by image sequence analysis, *SPIE-IS&T Electric Imaging*, Vol.5013, pp.186-193, Santa Clara, USA.

Lucas, B., D. and Kanade, T., 1981 An iterative image registration technique with an stereo vision, *DARPA Image Understanding Workshop*, pp.121-130, Washington D.C., USA



(a) View from left side



(b) View from right side

Figure 7. 3D model of urban area