

EFFICIENT LINE EXTRACTION FOR DIGITAL ARCHIVES OF CULTURAL HERITAGE USING OPTICAL FLOW AND TRIFOCAL TENSOR

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 V, WG V/4

KEY WORDS: Cultural Heritage, Extraction, Sequences, Matching, Modelling, Visualization

ABSTRACT:

In order to perform object modelling by image processing procedures, line or feature extraction and stereo matching used to perform. However, there are some issues for efficient 3D modelling of objects. In particular, efficient line matching for reconstruction of objects is needed to be resolved. With this objective, a robust line matching method which includes optical flow and trifocal tensor has been investigated. This paper focuses on more efficient line matching using Least-Median Squares (LMedS) and automatic 3D modelling of historical structure.

1. INTRODUCTION

Recently, efficient spatial data acquisition and visualization have been receiving more attention from the view point of construction of digital city, VR museum including digital archives. Generally, in order to perform object modelling through digital images, line or feature extraction and stereo matching are performed, and various matching methods such as area based matching, future based matching have been proposed. In particular, line often gives important information for object extraction, and accurate 3D results depend on rigorous line extraction and matching.

With this motive, the authors have been concentrating on developing an efficient line matching procedure for 3D modelling of historical structure using image sequences by video camera. The 3D modelling method contains line matching which is performed by line extraction and line tracking. The line extraction was performed by Canny operator (Canny, 1986), and line matching was performed using optical flow and trifocal tensor (Beardsley, et al., 1996). Therefore, availability of image sequences for object modelling was demonstrated (Kunii and Chikatsu, 2003). However, there are still some issues for 3D modelling of historical structure such as acquisition of image sequences or automatic line extraction for the first image, and these issues are needed to be resolved for more convenient and efficient 3D modelling.

In these circumstances, this paper focuses on more efficient line matching method using Least-Median Squares (LMedS). The LMedS have ability to calculate the trifocal tensor more accurately than least squares method, and more accurate line matching can be performed. In addition, the LMedS is able to remove inaccurate matched lines during the line matching procedure.

This paper describes more efficient line matching method using LMedS, and investigates an adaptability of this system to 3D modelling of historical structures "Koma house" which was built in 17th century (about 300 years ago), and designated as national important cultural assets in 1971.

2. LINE MATCHING BY OPTICAL FLOW

Image sequences for the Koma house were taken while amateur video camera was moving in horizontal direction by hand held, and line matching was performed by following procedures in Figure 1. Detail procedures of the line matching method are as follows.

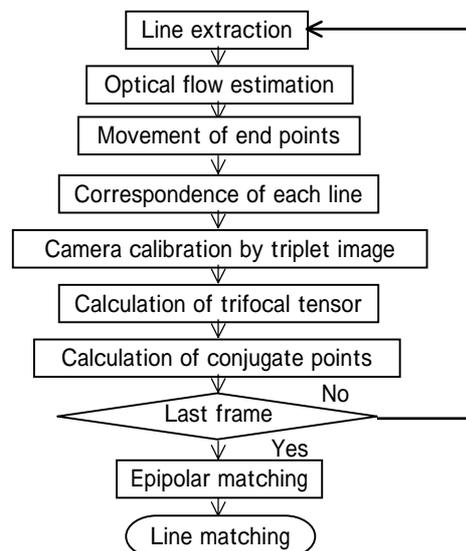


Figure 1. System flow of line matching

2.1 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 an interest point, and the reliability is an index for representing influence of noise. Therefore, the height h and the reliability r are calculated by following equation.

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

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

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 first image of the image sequences for Koma house, and Figure 3 shows the extracted lines by the method for the first image.



Figure 2. Koma house (the first image)



Figure 3. Line extraction

2.2 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 4.

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

where,

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

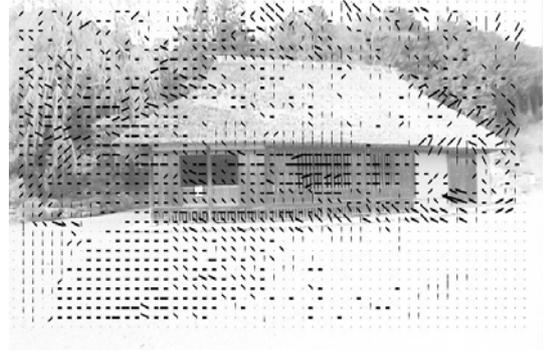


Figure 4. Optical flow estimation

2.3 Similarity Function

The line matching between each frame was performed by similarity function. The similarity is an index for representing how similar are the 2 lines. Figure 5 shows extracted 2 lines L_i ($n-1$ frame) and L_j (n frame), and L_m shows generated middle separating line. Let the average distances between L_m and 2 lines are D_i and D_j , and the projected lengths to the L_m are P_c and P_w , respectively. Thus, the similarity of 2 lines is calculated by following equation.

$$S = \frac{1}{k_d D + k_p P} \quad (4)$$

where,

$$D = D_i + D_j, P = \frac{P_w}{P_c} (P_w > P_c) \text{ or } P = \frac{P_c}{P_w} (P_c > P_w)$$

k_d, k_p : coefficient parameters ($k_d = 0.3, k_p = 0.7$)

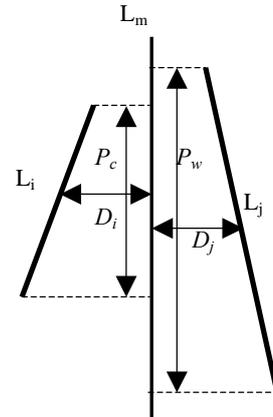


Figure 5. Similarity of 2 lines

In order to perform line matching efficiently, threshold value of the similarity was needed. The threshold value was calculated by theory of error diffusion using pointing accuracy of each line ends (± 1 pixel) and accuracy of the optical flow estimation (Ong

and Spann, 1999). Thus, the threshold value could be set automatically by following equation.

$$S_t = S_{\max} - \sigma_s \quad (5)$$

where,

$$\sigma_s^2 = \left(\frac{\partial S}{\partial D}\right)^2 \cdot \sigma_D^2 + \left(\frac{\partial S}{\partial P}\right)^2 \cdot \sigma_P^2$$

S_{\max} is maximum value of S , and σ_D and σ_P are accuracy for D and P . Therefore, S_t which is threshold value of S can be calculated automatically, and the line matching can be performed.

3. LINE MATCHING BY TRIFOCAL TENSOR

Some unmatched lines will remain in the above optical flow estimation procedures. In order to revive unmatched lines, trifocal tensor was adopted in this paper. Details of the line matching by trifocal tensor are as follows.

3.1 Trifocal Tensor

Trifocal tensor is geometric relation between 3 images which were taken from different camera positions for the same object. 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 common points to 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.

$$\begin{aligned} -z_2 z_3 g_{22} + z_2 y_3 g_{23} + y_2 z_3 g_{32} - y_2 y_3 g_{33} &= 0 \\ z_2 z_3 g_{21} - z_2 x_3 g_{23} - y_2 z_3 g_{31} + y_2 x_3 g_{33} &= 0 \\ z_2 z_3 g_{12} - z_2 y_3 g_{13} - x_2 z_3 g_{32} + x_2 y_3 g_{33} &= 0 \\ -z_2 z_3 g_{11} + z_2 x_3 g_{13} + x_2 z_3 g_{31} - x_2 x_3 g_{33} &= 0 \end{aligned} \quad (6)$$

where,

$$g_{ij} = x_i t_{1ij} + y_i t_{2ij} + z_i t_{3ij}$$

These 4 equations are generated by one common conjugated point to 3 images. The trifocal tensor has 27 ($=3 \times 3 \times 3$) unknown parameters which are able to calculate by more than the same number of equations. Therefore, more than 7 common conjugated points are needed to 3 images, and unmatched points in the third image are calculated by these above equations.

3.2 Least-Median Squares (LMedS)

In order to acquire unknown parameters by observation equations, calculation by least squares method is used. However, observation equations for the trifocal tensor are only acquired for common conjugated points to 3 images, and calculation of the trifocal tensor depend on the accuracy of the matched line. Therefore, calculation of the trifocal tensor by Least-Median Squares (LMedS) (Rousseeuw and Leroy, 1986) was investigated in this paper. The LMedS is one of the numerical calculation method by least squares using only accurate observation equations. Detail procedures of the LMedS are as follows:

(1) The trifocal tensor is calculated by these observation equations using least squares method, and observation

errors for each equation which are generated by least squares are acquired.

(2) A median value for these observation errors is extracted, and threshold value for the observation error $\hat{\sigma}$ is calculated by following equation.

$$\hat{\sigma} = C \left\{ 1 + \frac{5}{n-F} \right\} \text{med} \sqrt{\varepsilon_i^2} \quad (7)$$

where,

C : coefficient value ($=1.4826$)

n : number of observation equations

F : number of unknown parameters

ε_i : observation error

(3) These observation equations which have the observation error more than the threshold value are removed, and the trifocal tensor is calculated by least squares using out of those equations.

(4) Above procedures are iterated until observation errors for all equations become less than the threshold value, and final result is adopted as accurate trifocal tensor.

In addition, lines which were corresponded for the removed equations were also removed during the line matching procedure. Consequently, the trifocal tensor could be calculated accurately, and useless lines could be removed efficiently.

4. 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 ($\varphi_1, \kappa_1, \omega_2, \varphi_2, \kappa_2$) 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 by using these epipolar lines efficiently, least squares matching (LSM) method was adopted in this paper (Gruen, 1985).

As a result, line matching for 81 lines could be performed correctly. These matched lines constitute surface of Koma house, and out of those useless lines such as background or timber were removed automatically. Figure 6 shows the result of the line matching.

5. 3D MODELLING

The line information for 3D modelling can be acquired efficiently by the method in the previous chapter. However, each surface on the Koma house is needed to be recognized for 3D modelling. Therefore, surface recognition was performed by morphological opening procedure, and the extracted surfaces were conjugated with the matched lines in this paper (Kunii and Chikatsu, 2003). Figure 7 shows the result of the opening procedure for the first image.

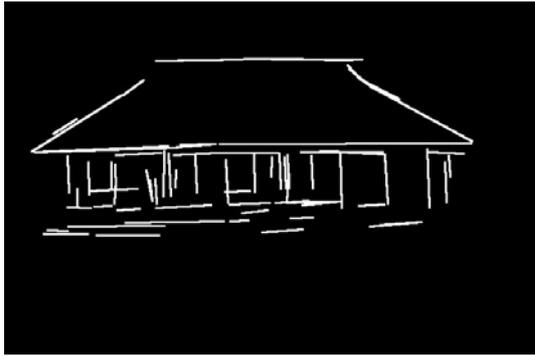


Figure 6. Result of line matching

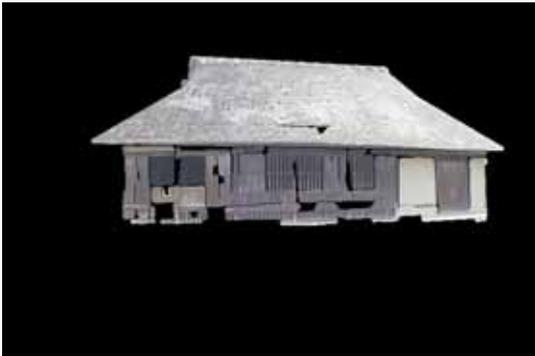


Figure 7. Result of surface extraction

Furthermore, in order to perform 3D modelling for the Koma house, camera calibration for the first image and the last image were performed by combined adjustment (Chikatsu and Kunii, 2002). Therefore, 3D data for the Koma house could be calculated efficiently. In addition, TIN for each surface was generated by using end points of the matched lines, and wire frame was reconstructed. In addition, texture mapping could be performed for each surface. Figure 8 shows wire frame model, and Figure 9 shows texture mapped model for the Koma house.

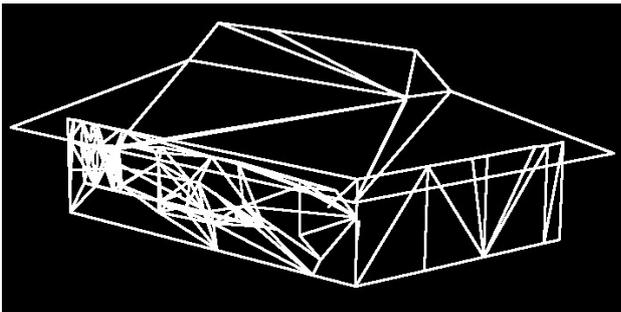


Figure 8. Wire frame model

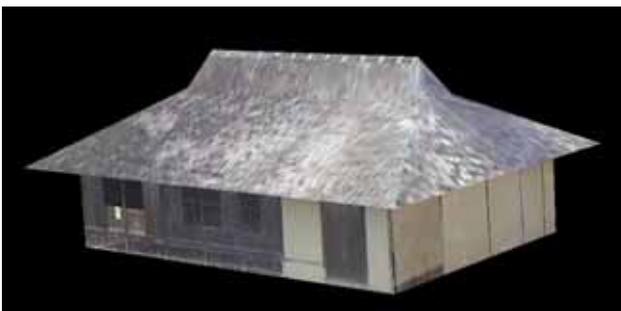


Figure 9. Texture mapped model

6. CONCLUSION

This paper investigates mainly 3 issues regarding 3D modelling for historical structure by using image sequences: (1) efficient and robust line matching method using optical flow, trifocal tensor and epipolar matching (2) calculation of accurate trifocal tensor, (3) efficient remove of the useless lines during the line matching procedure and followings main results were obtained:

- + Line matching was improved by similarity function with threshold value which acquired automatically.
- + LMedS realized improvement of trifocal tensor and remove of the useless lines.

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

- + Efficient texture mapping.
- + 3D modelling for more complicated object.

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.

Ong, E., P. and Spann, M., 1999. Robust optical flow computation based on Least-Median-of-Squares regression, *International Journal of Computer Vision*, Vol.31, No.1, pp.51-82.

References from Books:

Rousseeuw, P., J. and Leroy, A., M., 1986. *Robust Regression and Outlier Detection*, Wiley

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. Efficient 3D modeling for historical structure by image sequences analysis. *International Archives of Photogrammetry and Remote Sensing*, Vol.XXXIV, Part5/W12, pp.201-204, Ancona, Italy.

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