Journal of the Korean Society of Remote Sensing, Vol. 11, No.3, 1995

Height Estimation of the Flat-Rooftop Strúctures using Line-Based Stereo Matching

Sung-Han Choi, Gi-Mun Um, Kwae-Hi Lee
Robotics Lab. E.E. Dep. Sogang Univ.

직선 기반 스테레오 정합을 이용한 평면 지붕 인공지물의 고도 정보 추출

최성한 · 엄기문 · 이쾌희
서강대학교 전자공학과 로보트 연구실

Abstract

In this paper, the algorithm to extract the height of flat-rooftop structures in stereo aerial image is suggested with an assumption that the location, orientation, focal length, and field of view of a camera are known. It can be adapted to stereo aerial or satellite images. For performing feature-based stereo matching, the line segments suitable to describe the shape of general buildings are chosen as the feature. This paper is composed of three categories; the first step is to extract edges of structures with the polygon extraction algorithm which utilizes the edge following method, the second step is to perform the line segment matching with the camera information, and the last step is to calculate the location of each matched line and to estimate heights. The stereo images used in experiments are not real but synthetic ones. The experiment shows good results.

요 약

본 논문에서는 카메라의 정확한 정보(카메라 자세, 위치, 초점거리, 시야각)를 알고 있다는 가정하에 스테레오 항공영상에서 평면지붕 인공지물의 고도정보를 추출하기 위한 알고리즘이 제안하였다. 이 알고리즘은 위성영상에 대해서도 적용이 가능하다. 특징기반 정합을 수행하는데 있어, 일반적인 건물을

Received 17 November 1995 ; Accepted 11 December 1995.

묘사하기 위해 적합한 직선을 특징으로 선택하였다. 이 논문은 세단계로 구성되어 있다. 첫 번째 단계에서는 에지 추적 기법을 이용한 다각형 인공지물의 경계 추출 알고리듬을 통하여 영상상에서 인공 지물의 경계를 추출한다. 두 번째 단계에서는 카메라 정보를 이용하여 인공지물의 구성하고 있는 직선성분들을 정합한다. 마지막 단계로 정합된 선분쌍을 이용하여 실제높이를 계산하고 높이 정보를 추출한다.

실험에 사용된 영상은 실제로 취득한 영상이 아닌 합성영상을 사용하였다. 실험 결과 본 논문에서 제안한 인공지물 경계 추출, 고도 정보 추출은 좋은 성능을 보임을 알 수 있었다.

1. Introduction

Unlike the natural structures, the man-made structures have discontinuity of heights(Shinha 1991). Therefore in extracting the height of buildings, the feature-based stereo matching algorithm is usually used rather than the aerial-based stereo matching algorithm(Medioni & Nevita 1985, Barnard & Tompson 1980). Line segments are the best features to describe shape of buildings in aerial images(Huertas & Nevatia 1988). And those are very useful in the feature-based stereo matching.

For aerial images, disparities are very small because the locations of the cameras are very high. So, in order to estimate the exact height of buildings, the accuracy of line parameters must be ensured.

The method of disparity calculation by pixel matching is apt to be affected by quantization error. This paper shows the way to extract height information directly from line-segment parameters of stereo image. The algorithm can be applicable to stereo aerial images, in condition that buildings have flat-rooftops, and camera information is known.

2. Border Line Extraction of Structures

The typical algorithms(Huertas & Nevatia 1988, Liow & Pavlidis 1990, Irvin & McKeown 1990) that find the edge of buildings are usually consist of several steps; edge filtering, edge selection, thinning edge lines, and calculating their parameters. These produce errors by smoothing in edge filtering, and also have a lot of complexity. So any algorithms have not given good results. Therefore, the extraction algorithm of the exact building contours(Choi & Lee 1995) which gives line parameters of edge lines directly from images, is used in this paper. It extracts low frequency areas, selects candidate regions among areas, and then gets the parameters of the building's line edges with edge following method(Sauer 1991). It has two benefits. One is a robustness against the noise. The other an accuracy of line parameters is ensured.

3. Building and Line Segment Matching

Since the location and orientation of the camera are assumed to be known, the real location on the ground corresponding to the point on a image can be estimated roughly. In this calculation, a perspective camera model is used. The estimated location has some error because the height is not known. The real locations on the ground are used to match buildings and line segments of a building.

Let (x,y) represents the point on the image coordinates, and let (X,Y,Z) represents a corresponding point on the world coordinates. The relationship between the image and world coordinates is given by eq (1).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = sR \begin{bmatrix} x \\ y \\ z \end{bmatrix} + P \begin{bmatrix} 0 \\ 0 \\ -f \end{bmatrix} \quad \cdots\cdots (1)$$

$$R = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & 0 & - \\ 0 & 1 & \\ \sin\gamma & & c \end{bmatrix} \quad \cdots\cdots (2)$$

$$(\alpha = roll,\quad \beta = p\,i\,tch,\quad \gamma = yaw)$$

Here, s is a scaling factor, R is a rotation matrix, C is a location of the camera focus, and f is a focal length. The real position on the ground corresponding to the point of the image coordinates lies in the straight line drawing from C to (X,Y,Z) point. A rough location can be acquired on the condition that the ground surface has only little height variation, and the average height of the ground is known.

In this paper, first, the matching of the buildings is performed with the length of the border on the ground and a weight center of corner points on the ground, and then matching of the lines is performed with orientations of the corner points.

The process of the structures matching is as follows.

① Calculate locations of coner points of structures' rooftop in image coordinates
② Transform coner points from image coordinate to world coordinates
③ Predict real locations of coner points on the ground
④ Calculate weight center of coner points about each structure
⑤ Calculate border line length
⑥ Match the structures in stereo image with weight centers
⑦ Confirm with border line length

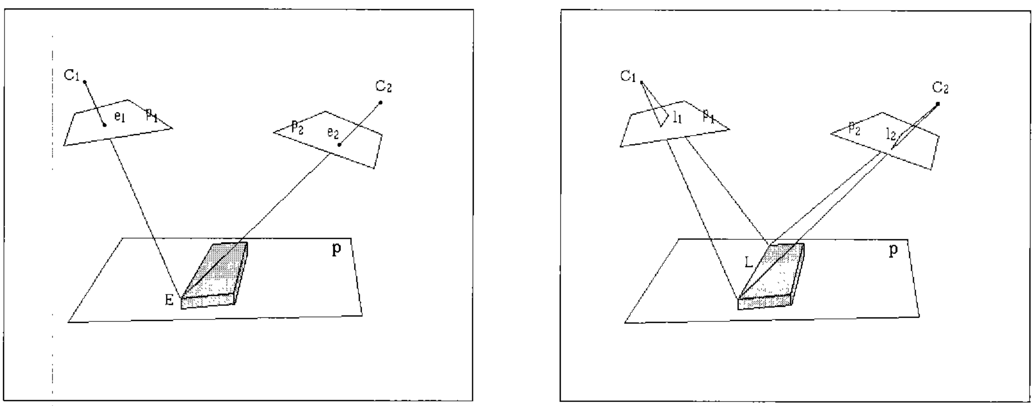
From this process, matched structures are found. And then, the line matching is performed with each matched structure

① Get vectors that the start point is weight center of coner point and the end point is coner point.

② Match the lines with orientation and length of vectors

4. Height Estimation

After the matching step, matched line pairs are acquired. One matched line pair comes from some edge line of the building.



(a) point-based height estimation (b) line-based height estimation

Fig. 1 Schematic Diagrams of Building's Height Extraction

Fig. 1 represents schematic diagrams which show the estimation of height with features. Fig 1-(a) show that a height information is extracted from the matched points(Sinha 1991). Theoretically, two lines $\overline{C_1e_1}$ and $\overline{C_2e_2}$ intersect on the point E, but practically, they are not intersect by quantization errors. Therefore a center of the line segment which makes a minimum distance between two lines, is regarded as an intersection point on the ground(Ghosh 1979). This method is usually used in the stereo vision system. But this method is affected by a quantization error.

Fig. 1-(b) shows that a height information is extracted from the matched line segments developed by Zhang(1994) and Schenk(1992). From the plane P_i including C_i and l_i (i=1,2), we can calculate the cross line of the two planes. The procedure is as follows.

First, line equations l_i of the image coordinates are transformed into L_i of the world coordinates by eq. (3).

$$L_i = R_i \left(l_i + \begin{bmatrix} 0 \\ 0 \\ -f_i \end{bmatrix} \right) + C_i \quad \text{(3)}$$

where,

R_i : rotation transformation, f_i : *focal length*

After getting L_i, we can calculate plane equations P_i including line L_i, and point C_i by eq. (4).

$$P_i = t_i L_i + (1 - t_i) C_i \quad \text{(4)}$$

From P_1 and P_2, we can calculate the cross line hat$\hat{L}$.

$$\hat{L} = P_1 \cap P_2 \quad \text{(5)}$$

Now we can extract the height information from the line segment $\hat{L}$. If the two planes are almost parallel to each other, the error of the estimated height will be large, and this case will be excluded. The number of the height informations that we can get will be at least 4 for each building because the most building's rooftop consist of more than 4 lines. In this paper, the height of buildings will be determined by an arithmetic average.

The merit of this method reduce an effect of a quantization error. The line equation of a line on a image is acquired by least mean square method. If a line is sufficiently long, we get a exact line equation from quantized points. Therefore accuracy of a estimated height from the matched lines is increased.

5. Experiments

1) Experimental Images

In Fig. 2 there is a perspective view including several structures to intend to extract height information on a plane. Each structures have flat-rooftop and some different value of height. Also Fig. 2 represents the coordinate system used in experiments. X-Y plane is the bottom plane.

Table 1 represents the real height of each structure.

Images used in experiments is acquired 3 different position and different orientations as

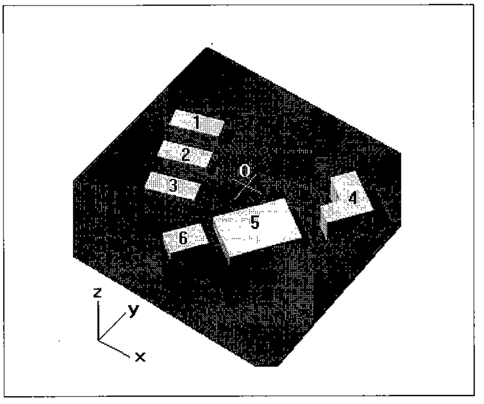


Fig. 2 Shape and Location of the Structures

Table 1 Real Heights of the Structures

# of Structure	1	2	3	4	5	6
Real Height	40.0	30.0	20.0	70.0	40.0	30.0

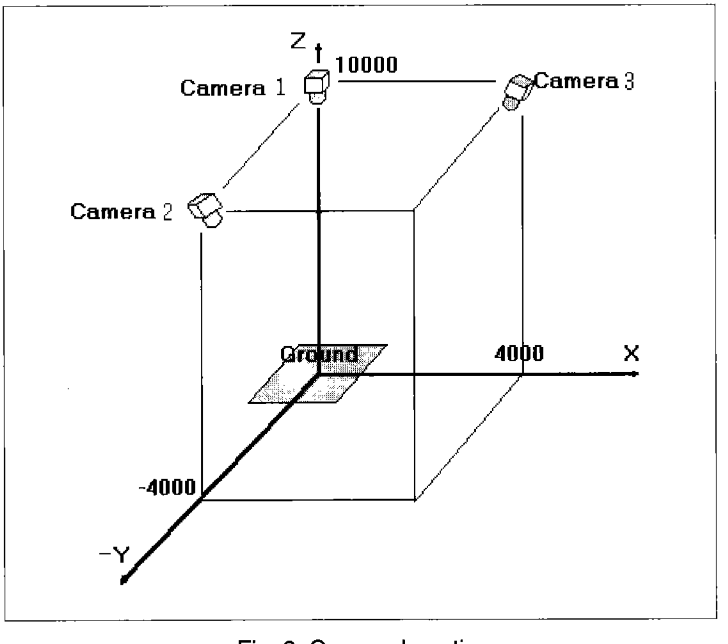
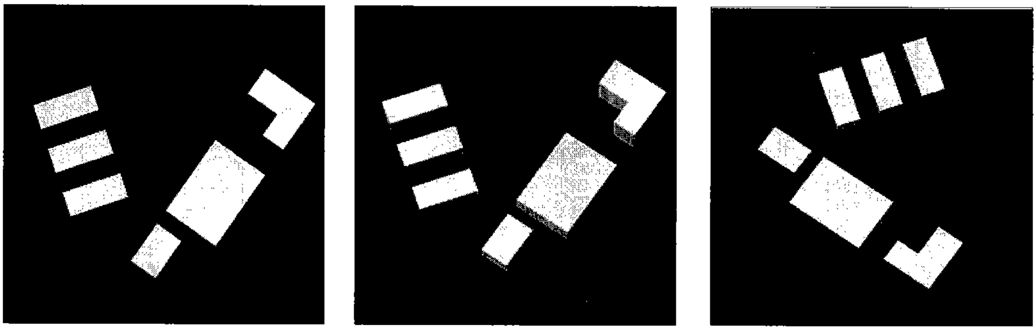


Fig. 3 Camera Location



(a) Image 1 (b) Image 2 (c) Image 3

Fig. 4 Experimental Images

Fig. 3. The height of cameras are the same, and all cameras are oriented toward origin of coordinate system. Fig. 4 shows 3 images for experiments. These images are not real but synthetic ones. The model used in making synthetic images is a perspective model that is usually used in camera modeling

Table 2 Camera & Image Specification

	Image 1	Image 2	Image 3
Camera Location (x,y,z)	(0, 0, 10000)	(0,-4000, 10000)	(4000, 0, 10000)
Camera Orientation (roll,pitch,yaw)	(0°, 0°, 0°)	(0°, 21.8014°, 0°)	(0°, 0°,-21.8014°)
Focal Length(cm)	0.4	0.4	0.4
Field of View (degree)	7.3333°	7.3333°	7.3333°
Image Size (pixel)	256 × 256	256 × 256	256 × 256
Length per Pixel(cm)	2.3913		

The specification of the cameras and images is on a Table 2.

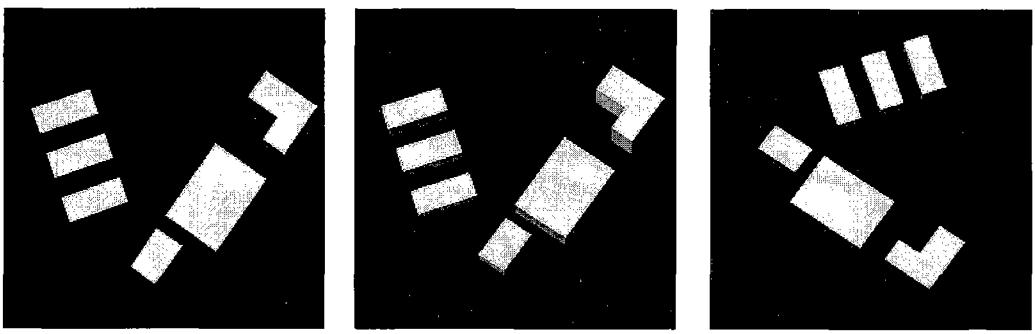
Table 3 Composition Stereo Image Pair

Stereo Pair 1	Image 1 + Image 2
Stereo Pair 2	Image 1 + Image 3
Stereo Pair 3	Image 2 + Image 3

Using this image, we make up of 3 stereo image pairs like Table 3.

2) Results

In this paper, we experiment in 3 stereo image pairs with the two method, line-based stereo matching and point-based stereo matching. The algorithm used in extracting of border line is the extraction algorithm of the exact building contours(Choi & Lee 1995) and it is commonly used in



(a) Image 1　　(b) Image 2　　(c) Image 3

Fig. 5 Extraction of Structures' Border Lines

two methods. Fig. 5 shows the results of the building edge extraction. The edges of building's roof-top regions are exactly acquired.

Table 4, 5, and 6 are represent the extracted height information and error rates.

Because the locations of cameras are higher than the locations of structures' rooftop, extracted heights are varied sensitively as variations of line parameters. In case of structure 1, 2, 3 in the stereo pair 2(Table 5), among the 4 border lines, two long lines are excluded because these are almost parallel with epipolar line. Two short lines' parameters are not exact because we recognize the side plane of structure as a rooftop plane. Therefore their error rates are higher than the others.

Table 4 Result of Stereo Pair 1

# of Structure		1	2	3	4	5	6
Real Height		40	30	20	70	40	30
Line based	Estimated	40.05	29.84	19.51	71.33	40.47	30.68
	Error(%)	0.1	0.4	2.5	1.9	1.2	2.3
Point based	Estimated	38.84	29.44	19.44	70.94	40.27	29.02
	Error(%)	2.9	1.9	2.8	1.3	0.7	3.3

Table 5 Result of Stereo Pair 2

# of Structure		1	2	3	4	5	6
Real Height		40	30	20	70	40	30
Line based	Estimated	37.87	27.39	18.20	69.15	39.60	29.52
	Error(%)	5.3	8.7	9.0	1.2	1.0	1.6
Point based	Estimated	37.75	27.29	16.89	69.18	39.51	27.76
	Error(%)	5.6	9.0	15.6	1.2	1.2	7.5

Table 6 Result of Stereo Pair 3

# of Structure		1	2	3	4	5	6
Real Height		40	30	20	70	40	30
Line based	Estimated	38.77	28.64	19.58	68.57	39.03	29.55
	Error(%)	3.1	4.5	2.1	2.0	2.4	1.5
Point based	Estimated	37.31	26.93	18.13	70.33	39.38	28.51
	Error(%)	6.7	10.2	9.4	0.5	1.6	5.0

Table 7 Average of Error rates

Stereo Pair	1	2	3
Line-based	1.40	4.46	2.60
Pixel-based	2.15	6.68	5.57

Table 7 shows average error rates of two methods. The result of the line-based stereo matching method is better than that of the point-based stereo matching method.

6. Conclusion

In this paper, the line edge parameters of the buildings are obtained by the extraction algorithm of the exact building contours, the matching of the buildings and lines are performed. Finally, the heights of the buildings are estimated from the matched line pairs with the camera information and matching results in two ways, the point-based and the line-based method. We had got good results by using an exact information of the camera. The result of the line-based method is better than that of the point-based method.

For the future work, we will apply this algorithm to real stereo image, and analyze an influence by camera information error.

Acknowledgment

This research has been supported by Agency for Defence Development in Korea.

References

A. Huertas and R. Nevatia, "Detecting Buildings in Aerial Image," *Computer Vision, Graphics, and Image Processing* Vol. 41, pp. 131-152, 1988.

Choi Sung-han, Lee Kwae-hee, "The Extraction of Exact Building Contours in Aerial Images", *Journal of KSRS* vol. 11, no 1, 1995.

G. Medioni, R. Nevita, "Segment-Based Stereo Matching," *Computer Vision, Graphics, and Image Processing,* Vol. 31, pp. 2-18, 1985.

Ken Sauer, "Enhancement of Low Bit Rate Coded Image Using Edge Detection & Estimation," *Graphical Models & Image Processing,* vol. 53, no 1, pp. 52-62, 1991.

R. Bruce Irvin and David M. McKeown Jr., "Toward Automatic Cartographic Feature Extraction," *Mapping and Spatial Modelling for Navigation,* NATO ASI Series, Vol.F65 pp.495-520, 1990.

Sarvajit S.Sinha, "Automatic Reconstruction of Buildings from Aerial Imagery," *SPIE Applications of Artificial Intelligence,* Vol.1468, pp.698-709, 1991.

S. K. Ghosh, *Analytical Photogrammetry,* Pergamon Press, 1979.

S. T. Barnard and W. B. Tompson, "Disparity Analysis of Image," IEEE Trans. *Pattern Analysis and Machine Intelligence,* Vol. PAMI-2, No. 4, pp. 333-340, 1980.

W. Cho, T. Schenk and M. Madani, "Resampling Digital Imagery to Epipolar Geometry," *Proc. of ISPRS,* Vol. 29 Part B3, Commission 3, pp.404-408, 1992.

Y. Liow and T. Pavlidis, "Use of Shadows for Extracting Buildings in Aerial Images," *Computer Vision, Graphics, and Image Processing* Vol. 49, pp. 242-277, 1990.

Zhengyou Zhang, "Estimating Motion and Structure from Correspondences of Line Segments Between Two Perspective Images," *INRIA Sophia-Antipolis n'2340,* September, 1994.