Image segmentation and line segment extraction for 3-d building reconstruction

Chul-Soo Ye*, Kyoung-Ok Kim*, Jong-Hun Lee*, Kwae-Hi Lee**

*Spatial Imagery Information Research Team,

Spatial Information Technology Center, ETRI- Computer & Software Research Laboratory

161 Gajeong-dong, Yuseong-gu, Daejeon, 305-600, Korea

E-mail : {csye, kokim}@etri.re.kr

**Dept. of Electronic Engineering, Sogang University

1 Shinsu-Dong, Mapo-Gu, Seoul, Korea, 121-742

Email : khlee@sognag.ac.kr

Abstract

This paper presents a method for line segment extraction for 3-d building reconstruction. Building roofs are described as a set of planar polygonal patches, each of which is extracted by watershed-based image segmentation, line segment matching and coplanar grouping. Coplanar grouping and polygonal patch formation are performed per region by selecting 3-d line segments that are matched using epipolar geometry and flight information. The algorithm has been applied to high resolution aerial images and the results show accurate 3-d building reconstruction.

Key words: image segmentation, 3-d building reconstruction, coplanar grouping.

1. Introduction

There have been many building detection and extraction techniques. In the last few years, the analysis of 3-D data such as 3-D lines, 3-D corners, or digital elevation model (DEM) has been much developed. Some methods based on coplanar grouping of 3-D lines have been proposed for reconstructing generic models of buildings from high-resolution multiple imagery (Bignone, 1995; Frere *et al*, 1997; Henricsson and Baltsavia, 1997; Ye and Lee, 2000). We propose a method to extract a building roof consisting of a set of planar polygonal patches by watershed segmentation and 3-D grouping.

2. Image segmentation using watershed

El-Fallah and Gary E. Ford represented the image as a surface and proved that setting the inhomogeneous diffusion coefficient equal to the inverse of the magnitude of the surface normal results in surface evolving speed that is proportional to the mean curvature of the image surface (El-Fallah and Ford, 1997). This model has the advantage of having the mean curvature diffusion (MCD) render invariant magnitude, thereby preserving structure and locality. By coupling the min/max flow to the surface diffusion model controlled by the surface's normal magnitude and smoothness, noise is eliminated and thin edges are preserved more efficiently (Ye and Lee, 2001).

After edge preserving filtering using mean curvature diffusion, a hybrid image segmentation algorithm is applied, which integrates edge-based and region-based techniques through the watershed algorithm (Ye and Lee, 2002). First, Images are segmented by watershed

algorithm, the segmented regions are combined with neighbor regions. Region adjacency graph (RAG) is employed to analyze the relationship among the segmented regions. The graph nodes and edge costs in RAG correspond to segmented regions and dissimilarities between two adjacent regions respectively. After the most similar pair of regions is determined by searching minimum cost RAG edge, regions are merged and the RAG is updated. The proposed method efficiently reduces noise and provides one-pixel wide, closed contours.

3. Line segment extraction from a region boundary

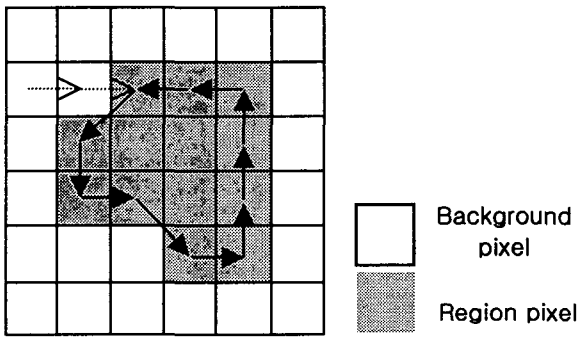


Fig. 1. Boundary represented by 8-directional chain code.

For each region in the labeled image, an 8-directional chain code is generated by following the boundary of the region in a counterclockwise direction and assigning a direction to each pixel as shown in Fig. 1.

We need to calculate discrete curvature of the boundary chain code to find control points used to detect line segments. Curvature for a function $y(x)$ is defined as

$$\kappa(x) = \frac{|\, y''(x)\,|}{(1+(y'(x))^2)^{3/2}} \tag{1}$$

where

$$y'(x) = \frac{dy}{dx}, y''(x) = \frac{d^2y}{dx^2}. \tag{2}$$

Because the curves taken from real images do not have a proper function $y = y(x)$, we cast x and y into parameter form, $x(t)$ and $y(t)$, where $0 < t < L$, L being the length of the curve. Then

$$\kappa(t) = \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}} \tag{3}$$

where

$$\begin{aligned} \dot{x}(t) &= \frac{dx}{dt},\ \ddot{x}(t) = \frac{d^2x}{dt^2}, \\ \dot{y}(t) &= \frac{dy}{dt},\ \ddot{y}(t) = \frac{d^2y}{dt^2}. \end{aligned} \tag{4}$$

Differentiating discrete signals is inherently noisy. Therefore we convolve $x(t)$ and $y(t)$ with the Gaussian kernel to smooth and differentiate the functions as follows:

$$\begin{aligned} X(t,\sigma) &= x(t) * g(t,\sigma) \\ Y(t,\sigma) &= y(t) * g(t,\sigma) \end{aligned} \tag{5}$$

where

$$g(t,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}. \tag{6}$$

We get the following formula for discrete curvature:

$$\kappa(t,\sigma) = \frac{\dot{X}(t,\sigma)\ddot{Y}(t,\sigma) - \ddot{X}(t,\sigma)\dot{Y}(t,\sigma)}{(\dot{X}^2(t,\sigma) + \dot{Y}^2(t,\sigma))^{3/2}}. \tag{7}$$

4. Line segment merging and position correction

After line segment extraction, we link line segments satisfying some criteria. For the two line segments AB and CD having similar direction as shown in Fig. 2, new line segment is created if the three conditions are satisfied.

$$|\theta_1 - \theta_2| < T_{\text{degree}} \tag{8}$$

$$\frac{AB+CD}{AD} > T_{length} \tag{9}$$

$$d_1 < T_{dist} \text{ and } d_2 < T_{dist} \tag{10}$$

For the small line segment BC between long line

segments, if the condition (11) is satisfied, find the intersection point H and create new line segments AH and HD as shown in Fig. 3.

$$AB + BC + CD > AH + HD \tag{11}$$

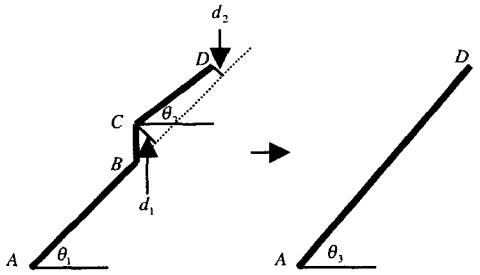


Fig.2. Merging two lines with similar direction.

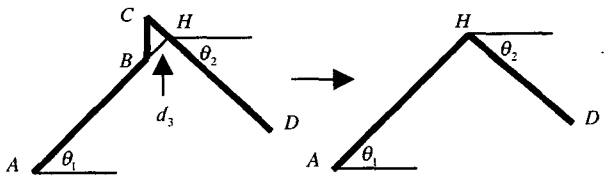


Fig. 3. Merging two lines with different direction.

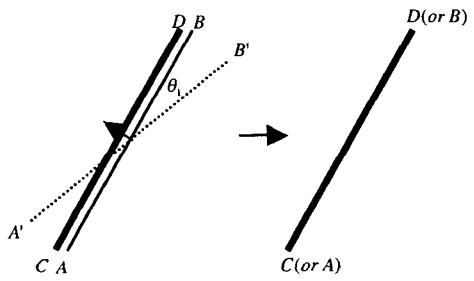


Fig. 4. Correction of line segment position.

To minimize the position error of the line segment extracted from a region boundary, the line segment is adjusted using edgeness measure function (Fig. 4). Discretize the domain of θ_1 between θ_{min} and θ_{max} in steps of $0.5°$. For every position in the 5×5 window located in the center of the line segment AB, compute the edgeness measure for θ between θ_{min} and θ_{max}. Find the location and θ with the maximum edgeness.

As the region in the roof is relatively large and its average segment length is long, we accept the region that satisfies the following 3 criteria:

- Total length of region boundary is longer than T_L.
- Number of control points is larger than T_C.
- Average segment length between control points is longer than T_A.

5. Segment stereo matching and 3-d grouping

Line segments extracted from the region boundary are matched in the other image by maximizing an edginess measure along epipolar line. The edginess measure is a function of the gradient magnitude and the local orientation of the gradient in the other image. For a maximum of the edginess measure, there is a high probability that there is an edge in the image. The epipolar constraint is applied to restrict the search space of the correspondences along the epipolar line.

The 3-D line segments obtained in stereo matching are to be grouped into coplanar configurations. The coplanar grouping process can be restricted to the 3-D line segments from region boundary. Coplanar grouping and polygonal patch formation is performed per region by selecting 3-D line segments that are matched using epipolar geometry and flight information. Starting with the longest ones, 2 line segments are selected in the region. If orthogonal distance from the line segments to a plane described by these line segments is small, we construct the plane that fits the line segments in a least-square sense. Coplanarity is assumed if both end points of the line segment have a distance to the plane smaller a particular bound and if the angle between the line segment and the plane's normal is close to $90°$ (Frere *et al*, 1997). All segments that satisfy this constraint are

then included in the defining set of the plane and the plane's equation is updated.

6. Experimental results

We choose the residential scene from the Avenches data set having the following characteristics: 1:5,000 image scale, vertical aerial photography, four-way image coverage, color imagery of size 1800×1800 pixels, ground area of approximately 75×75 millimeters (Fig. 5). An example of the image of size 400×400 from the Avenches data set is shown in Fig. 6(a).

Fig. 5. An image from the Avenches data set.

We see that noise within region is well removed by the proposed mean curvature diffusion (Fig. 6(c) and (d)). It is not necessary to merge regions into one roof plane as like other segmentation method (Fig. 6(e) and (f)). Fig. 7 shows examples of extracted roof boundary. Final result of line segment merging and position correction is shown in Fig. 8. Matched line segments after the stereo matching for the roofs of the building are shown in Fig. 9. Another result of roof reconstruction is shown in Fig. 10. Line segments were merged correctly from partially incorrect segmentation result as shown in Fig. 10 (c) and (d). The accuracy is evaluated with the reconstructed plane and the corner points in a CAD model of the building extracted with an accuracy of about 0.1 meters. Standard deviation of the coordinate difference between reference corner points and reconstructed corner points were computed including the rms-values for the distance between the reconstructed plane and the corner points of reference plane in a CAD model of the building in table 1 and table 2. The reconstructed plane was generated using the 3-d line segments satisfying the coplanarity.

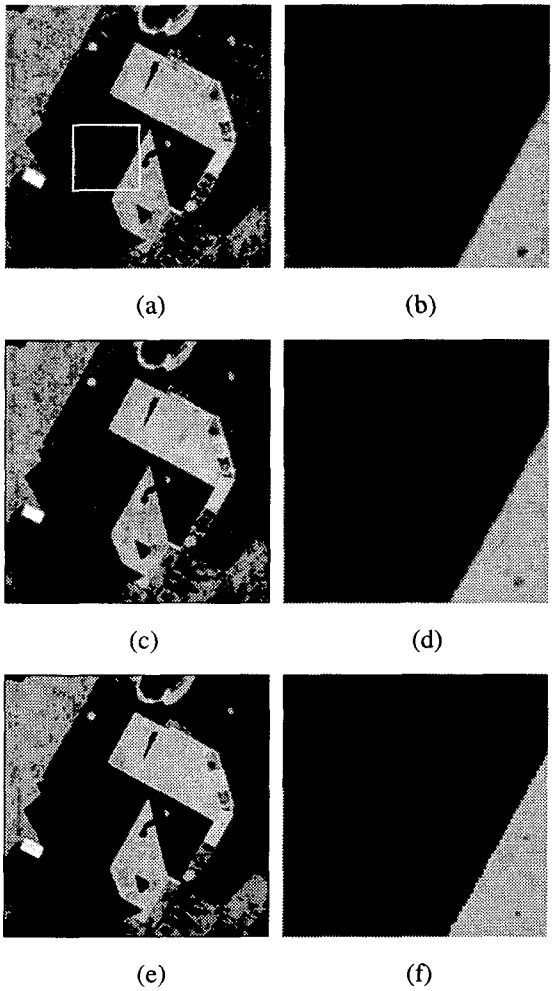
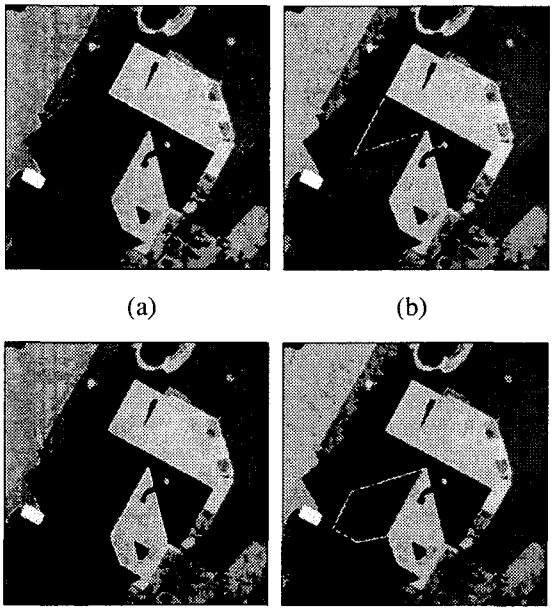


Fig. 6. (a) and (b) An example of aerial image from the Avenches data set and zoomed region of Fig. 6(a) (surrounded by the white box) (c) and (d) mean curvature diffusion filtered images (e) and (f) final segmentation results.



(a) (b)

(c) (d)

Fig. 7. Examples of extracted roof boundaries.

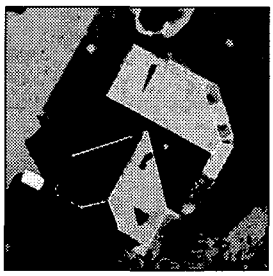


Fig. 8. Result of line segment merging and position correction.

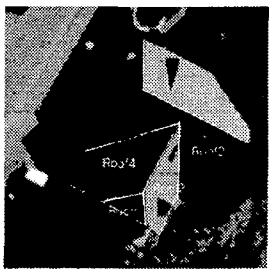
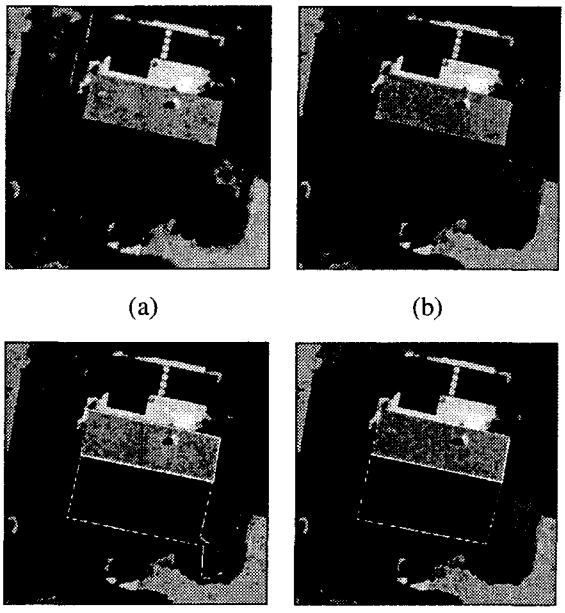
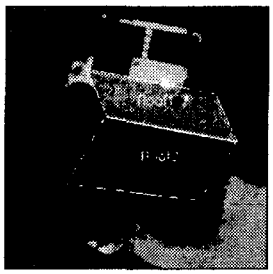


Fig. 9. Matched line segments after stereo matching for the building roof in Fig. 8 (matches in white).



(a) (b)

(c) (d)



(e)

Fig. 10. Another result of roof reconstruction (a) original image (b) segmented image (c) extracted roof boundaries (d) result of line segment merging and position correction (e) matched line segments.

Table 1. Accuracy achieved by coplanar grouping for the roofs in Fig. 9 [m].

	standard deviation of ΔX	standard deviation of ΔY	standard deviation of ΔZ	RMSE
Roof 1	0.0476	0.0143	0.0598	0.1190
Roof 2	0.0779	0.0508	0.2812	0.4783
Roof 3	0.0548	0.0509	0.1042	0.1999
Roof 4	0.0850	0.0799	0.2366	0.3032
Roof 5	0.0740	0.0529	0.0560	0.7701

Table 2. Accuracy achieved by coplanar grouping for the roofs in Fig. 10 (a)[m].

	standard deviation of ΔX	standard deviation of ΔY	standard deviation of ΔZ	RMSE
Roof 1	0.0224	0.1397	0.3064	0.2729
Roof 2	0.0499	0.0849	0.1718	0.2288

7. Conclusions

An approach for line segment extraction for 3-d building reconstruction was proposed. Building roofs are described as a set of planar polygonal patches, each of which is extracted by watershed-based image segmentation, line segment matching and coplanar grouping. We extracted line segments from a region boundary after image segmentation and then line segment merging was performed. To minimize the position error of the line segment extracted from a region boundary, the line segment was adjusted using edgeness measure function. Coplanar grouping and polygonal patch formation are performed per region by selecting 3-d line segments that are matched using epipolar geometry and flight information. The algorithm has been applied to aerial images and the results show correct line segment extraction from region boundary and accurate 3-d building reconstruction.

References

Bignone, F., 1995, Segment stereo matching and coplanar grouping, Technical Report BIWI-TR-165, Institute for Communications Technology, Image Science Lab, ETH, Zürich.

C.S. Ye and K.H. Lee, 2000, 3D building reconstruction using building model and segment measure function, *Journal of KITE*, vol. 37SP, no.4, pp. 46-55.

C.S. Ye and K.H. Lee, 2001, Anisotropic diffusion for building segmentation form aerial imagery, in *Proc. of International Symposium on Remote Sensing*, EMSEA and KSRS, Seogwipo, Korea, Oct. 31-Nov. 2, pp. 599-604.

C.S. Ye and K.H. Lee, 2002, Remote sensing image segmentation by a hybrid algorithm, *Journal of KSRS*, vol. 18, no.2, pp. 107-116.

El-Fallah, A.I., and Ford, G.E, 1997, Mean curvature evolution and surface area scaling in image filtering, *IEEE Trans. Image Process.*, vol. 6, no. 5, pp. 750-753.

Frere, D., M. Hendrickx, J. Vandekerckhove, T. Moons, and L. Van Gool, 1997, On the reconstruction of urban house roofs from aerial images, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser, Basel, pp.87-96.

Henricsson, O. and Emmanuel Baltsavias, 1997, 3-D building reconstruction with ARUBA: A qualitative and quantitative evaluation, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser, Basel, pp.65-76.