AUTOMATIC BUILDING EXTRACTION BASED ON MULTI-SOURCE DATA FUSION

Yi Hui Lu and John Trinder

School of Surveying and Spatial Information Systems
University of New South Wales, NSW 2052 Australia j.trinder@unsw.edu.au

KEY WORDS: Data Fusion, Image matching, Image understanding, Level Set formulation

ABSTRACT:

An automatic approach and strategy for extracting building information from aerial images using combined image analysis and interpretation techniques is described in this paper. A dense DSM is obtained by stereo image matching. Multi-band classification, DSM, texture segmentation and Normalised Difference Vegetation Index (NDVI) are used to reveal building interest areas. Then, based on the derived approximate building areas, a shape modelling algorithm based on the level set formulation of curve and surface motion has been used to precisely delineate the building boundaries. Data fusion, based on the Dempster-Shafer technique, is used to interpret simultaneously knowledge from several data sources of the same region, to find the intersection of propositions on extracted information derived from several datasets, together with their associated probabilities. A number of test areas, which include buildings with different sizes, shape and roof colour have been investigated. The tests are encouraging and demonstrate that the system is effective for building extraction, and the determination of more accurate elevations of the terrain surface.

1. INTRODUCTION

Stereo image matching determines corresponding pixels or features in two overlapping images and is the fundamental to digital photogrammetry for elevation determination. Conventional image matching techniques only supply a Digital Surface Model (DSM). This means that matching occurs on the top of man-made objects such as buildings, or on the top of the vegetation rather than the terrain surface and hence does not represent the terrain surface [Baltsavias et al 1995,Henricsson et al 1997 and Tönjes 1996]. The approach used in this research for 3D reconstruction from stereo images over trees or built-up areas is based on an attempt to understand and interpret the image content, and is significantly different from the current approaches for determining elevations on the terrain. In order to provide an accurate DTM of the terrain surface, the characteristics of the terrain cover, such as buildings and trees must to be determined. Although many automatic building extraction algorithms have been proposed by researchers (Collin et al 1998, Hanson et al 2001 and Henricsson 1998), there are no operational algorithms because each method is focused on a particular application area and is not transferable to other features. An important aspect of the proposed automatic building extraction system is that it aims to use image interpretation, as well as elevation information to extract the building areas by integrating image analysis and interpretation techniques.

2. GENERAL DESCRIPTION OF SYSTEM

Figure 1 illustrates the architecture of the automatic building extraction system. The goal of the component of the process is to extract accurate building boundaries for reconstruction of terrain elevations from overlapping aerial or satellite images over a variety of terrain types and ground cover.

The overall system consists of three main parts. Part 1 performs the matching of the stereo image pair, derives a disparity map, and produces a digital surface model (DSM). Then an analysis of the multispectral image supplies the results of multi-band classification, segmentation by classification and Normalised Difference Vegetation Index (NDVI). The four information layers shown as green in Figure 1 finally produce building interest areas. Part 2 uses information from part 1 to define an initial curve leading to the level set formulation of curve and surface motion to define the desired building boundaries, driven by an image-dependent speed function. Part 3 presents Dempster-Shafer fusion theory, which is used to combine different data sources to extract the correct building areas.

The DSM, obtained from LH Systems' Socet Set v4.2 is represented by a dense sample of points in order to avoid missing some structures. The derived DSM is then interpolated to the size of the original image for further processing.

Multispectral image classification is typically used to detect individual object primitives. It aids in reducing the complexity of the image content for the next processing step of feature detection. In order to find building areas, K-Means unsupervised classification is used to classify the image because it is a fully automatic process. then using a post classification procedure, a segmented image can be created from a classified image. Segmentation partitions a classified image into meaningful regions of connected pixels that are contained in the same class. The NDVI (Vegetation Index) can then be used to transform the multispectral data into a single image band representing vegetation. The NDVI (Normalized Difference Vegetation Index) values indicate the amount of green vegetation present in the pixel.

While multispectral images supply abundant information for land cover classification, the NDVI and DSM are two key parameters which define the difference between vegetated and non-vegetated objects. Simplistically, the areas which have heights above some limit, are likely to be either trees or buildings. Areas with low NDVI, and are above the general terrain surface are likely to be buildings, whereas areas with high NDVI and are above that surface are likely to be trees. Areas with high NDVI, with heights similar to the terrain surface are likely to be grassland or cultivated areas. Four information layers of the land cover classification, the results of the segmentation by K-means, DSM and NDVI, are input to ArcView Map Query operation to extract building interest areas.

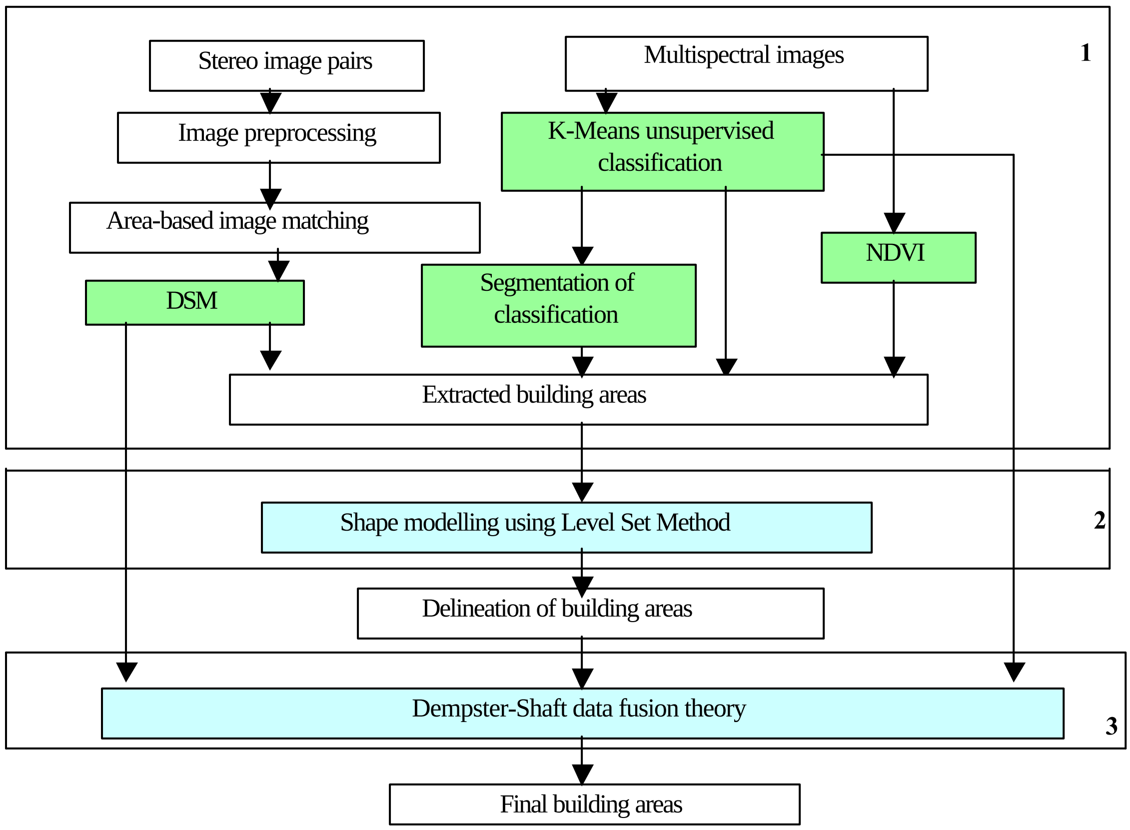

Figure 1 Architecture of the building extraction system

The level set method for curve propagating interfaces was introduced by Osher and Sethian (1988, 1999). It is based on mathematical and numerical work of curve and surface motion by Sethian (1985), and offers a highly robust and accurate method for tracking interfaces moving under complex motions. Based on the building interest areas derived from image analysis, the level set algorithm can be used to process all building interest areas to delineate their boundaries.

Finally, the Dempster-Shafer method is used to fuse three data sources: the DSM, the results of the multispectral classification, and the delineated building boundaries. Assume a set of n propositions making up the hypothesis space as denoted by Θ. 2^{Θ} are the subsets of Θ. Based on the information from the data sources, a probability mass **m** can be assigned to any proposition or union of propositions. For $\forall A \in 2^{\Theta}$. **m** is defined for every element A and the mass value $\mathbf{m}(A)$ is in the interval [0,1].

The following mass equations can be obtained:

$$\mathbf{m}(\varnothing)=0$$
$$\mathbf{m}(\Theta)=\sum_{A \subset 2^{\Theta}} \mathbf{m}(A)=1 \qquad (1)$$

where $\varnothing$ is the empty set.

In image classification, Θ is the set of hypotheses about a pixel class. The Dempster-Shafer theory permits the consideration of any subset of Θ. Applied to image classification problems, it means that not only single classes, but also any union of classes can be represented. The number of classes (including all possible unions, but excluding the null set) is called the power set and is equal to $2^n - 1$. For example, if n=3, $2^3 - 1 = 7$, classes are given by C1, C2, C3, C1 $\cup$ C2, C1 $\cup$ C3, C2 $\cup$ C3, and C1 $\cup$ C2 $\cup$ C3 (Klein 1999 , Hegarat-Mascle 1997 and Shafer 1976).

The Dempster-Shafer theory provides a representation of both imprecision and uncertainty through the definition of two parameters: support (Sup) and plausibility (Pls). They are obtained from the probability mass function **m**. Support for a given proposition means that all masses assigned directly by the data sources are summed. Plausibility for a given proposition means all masses not assigned to its negation are summed.

For multi-source evidential reasoning based region evaluation, the probability masses are assigned based on information provided by each image. This method is more reliable and is able to take into account a larger variety of situations.

3. TESTS AND RESULTS

While a number of test areas have been investigated, the results of processing one area will be given in the following. The test image is a pair of colour aerial images with 522 $\times$ 584 pixels in the row and column directions respectively, with GSD of 0.3 metre. The flying height is 3070 metres and the original scale of the images was 1:20,000. The scale of the images is smaller than desired, but larger scale images of the area were not available. The majority of buildings have white or red roofs, but there are some dark roof buildings as well. There is good contrast between the buildings and the background. Most of the building areas have been detected, but the dark roofed buildings are completely missed and will not be recovered. Some red roofed buildings are partly detected. Also, one car is assigned as a building in the top of image. Since only the colour image is available, the results of processing by Visible Vegetation Index (VVI) can be obtained. These areas with high VVI represent the vegetation and the areas with low VVI represent the ground and building areas. The Dempster-Shafer method provides a single value for the probability of the intersection (union) of the propositions. The consequence of the data fusion is that most incorrect building areas have been detected and deleted in the final result, as shown in Figure 2. There are 50 buildings in the image and 40 building are detected. The detection rate is 80%.

The part 1 of DSM extraction and low-level image analysis and interpretation in Figure 1 is important, since it supplies building interest areas and level set modelling and data fusion are based on these areas. As mentioned before, some dark roof buildings have been missed in the processing of part 1and will never be recovered. The results of building extraction can be improved by refining part 1 and using larger scale images. The improvement in building extraction by using the data fusion approach is 3 out of 50. This is important processing step because it makes the system to supply more reliable extracted buildings.

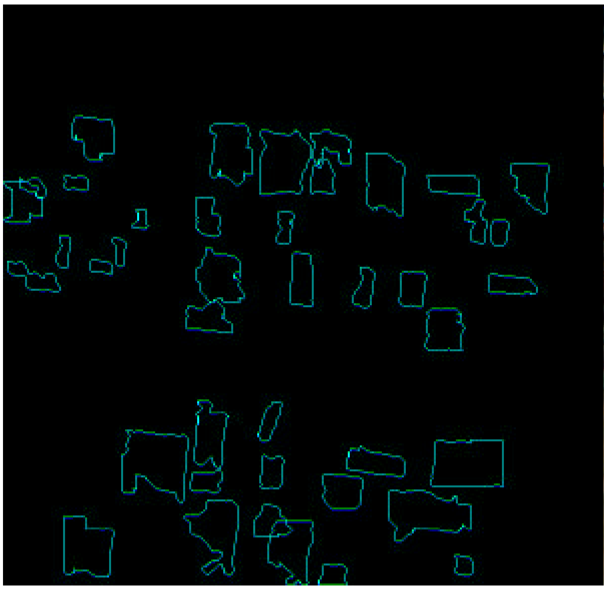

Figure 2 Result of Dempster-Shafer based on 3 datasets

Figure 3 Extracted buildings overlaid on ortho image

References:

Baltsavias E., Mason S. and Stallmann D. 1995. Use of DTMs/DSMs and Orthoimages to Support Building Extraction. *Automatic Extraction of Man-made Objects from Aerial and Space Images.* Birkhauser Verlag, Basel: 199-210.

Collin, R., Jaynes, C., Cheng, Y., Wang, X. and Stoller, F. (1998). The Ascender System: Automated Site Modeling from Multiple Aerial Images. *Computer Vision and Image Understanding* , 72(2):143-162.

Hanson, A., Marengoni, M., Schultz, H., Stolle, F., Riseman, E. and Jaynes, C. (2001). Ascender II: A Framework for Reconstruction of Scenes from Aerial Images. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Hegarat-Mascle S., Bloch I. and Vidal-Madjar D. 1997. Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (4): 1018-1031.

Henricsson O., (1998) The Role Color Attributes and Similarity Grouping in 3-D Building Reconstruction, *Computer Vision and Image Understanding*, 72(2): 163-184.

Henricsson O., Bignone F., Willuhn W., Ade F., and Kuebler O. 1997. Project AMOBE: Strategies, Current Status and Future Work. *International Archives of Photogrammetry and Remote Sensing,* 31(3):321-330.

Klein L. 1999. Sensor and Data Fusion Concepts and Applications. *SPIE Optical Engineering Press.*

Sethian J. A. 1985. Curvature and the evolution of fronts, *Comm. In Math. Phys*. 54: 487-499.

Osher S. and Sethian J.A. 1988. “Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulation”, *J. of Computational Physics*, 79: 12-49

Sethian J. A. 1985. Shape Modelling with front Propagation: A Level Set approach, *IEEE Transactions on Pattern analysis and Machine Intelligence,* 17(2):158-175.

Sethian J. A 1999. Level set methods and fast marching methods, *Cambridge University Press.*

Shafer G. 1976. A Mathematical Theory of Evidence. *Princeton University Press, Princeton, NJ.*

Tönjes R. 1996. Knowledge Based Modelling of Landscapes. *International Archives of Photogrammetry and Remote Sensing,* 31(3):868-873.