AUTOMATIC BUILDING EXTRACTION FOR 3D TERRAIN RECONSTRUCTION USING INTERPRETATION TECHNIQUES

Yi Hui Lu [1], **John Trinder** [1] **and Kurt Kubik** [2]

[1] School of Surveying and Spatial Information Systems
University of New South Wales, NSW 2052 Australia

[2] Department of Computer Science and Electrical Engineering
University of Queensland, QLD 4072 Australia

ABSTRACT

Conventional image matching techniques for DTM determination are unable overcome the disparity discontinuities in the stereo model caused by man-made structures and only supply a Digital Surface Model (DSM). In order to produce a DTM of the bare earth, the characteristics of the terrain cover, such as buildings and trees must to be determined to reduce the elevations derived from image matching to the terrain surface. An automatic approach and strategy for extracting building information from aerial images using combined image analysis and interpretation techniques is described in this paper. The approach is undertaken in three parts. In the first part, 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. In part two, based on the approximate building areas derived in part one, a shape modelling algorithm based on the level set formulation of curve and surface motion has been used to precisely delineate the building boundaries. Since the complex urban scene can result in wrongly extracting building regions, a method is required to evaluate the reliability of the evidence of buildings and delete those buildings that have been wrongly identified. Data fusion in remote sensing, such as the Dempster-Shafer approach, is undertaken to combine several image data sets for the purpose of information extraction. It can be used to interpret simultaneously knowledge from several data sources of the same region to find the intersection of propositions on extracted information derived from these datasets, together with their associated probabilities. In the third part, the Dempster-Shafer data fusion technique provides the theoretical basis for evaluating the reliability of the extracted buildings from the combination of the different data sources by a statistically-based classification. 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 three parts of the system are important procedures for effective building extraction, and the determination of more accurate elevations of the terrain surface.

INTRODUCTION

One of the most challenging problems in the fields of computer vision and digital photogrammetry is 3D reconstruction of the terrain surface from complex aerial images in urban or suburban areas where buildings, roads, trees and vegetation are intermingled in an intricate and complex fashion. Digital Terrain Models (DTM) produced by stereo image matching algorithms have been one of the primary goals of cartography for many years. Recently, interest in the area has been stimulated by the need for digital orthophotos, 3D city models, 3D building reconstruction, production and management of 3D databases for urban and town planning and Geographic Information Systems (GIS) modelling.

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.

Section 2 in this paper introduces the proposed system for automatic building extraction. Section 3 and Section 4 describe multispectral image analysis and the level set shape modelling process respectively. Dempster-Shafer data fusion theory is presented in Section 5. Section 6 gives experimental results, and conclusions are drawn in Section 7.

GENERAL DESCRIPTION OF AUTOMATIC EXTRACTION OF BUILDING SYSTEM

Figure 1 illustrates the architecture of the automatic building extraction system. Only the level set and Dempster-Shafter algorithms will be described in detail in this paper. The goal of these 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.

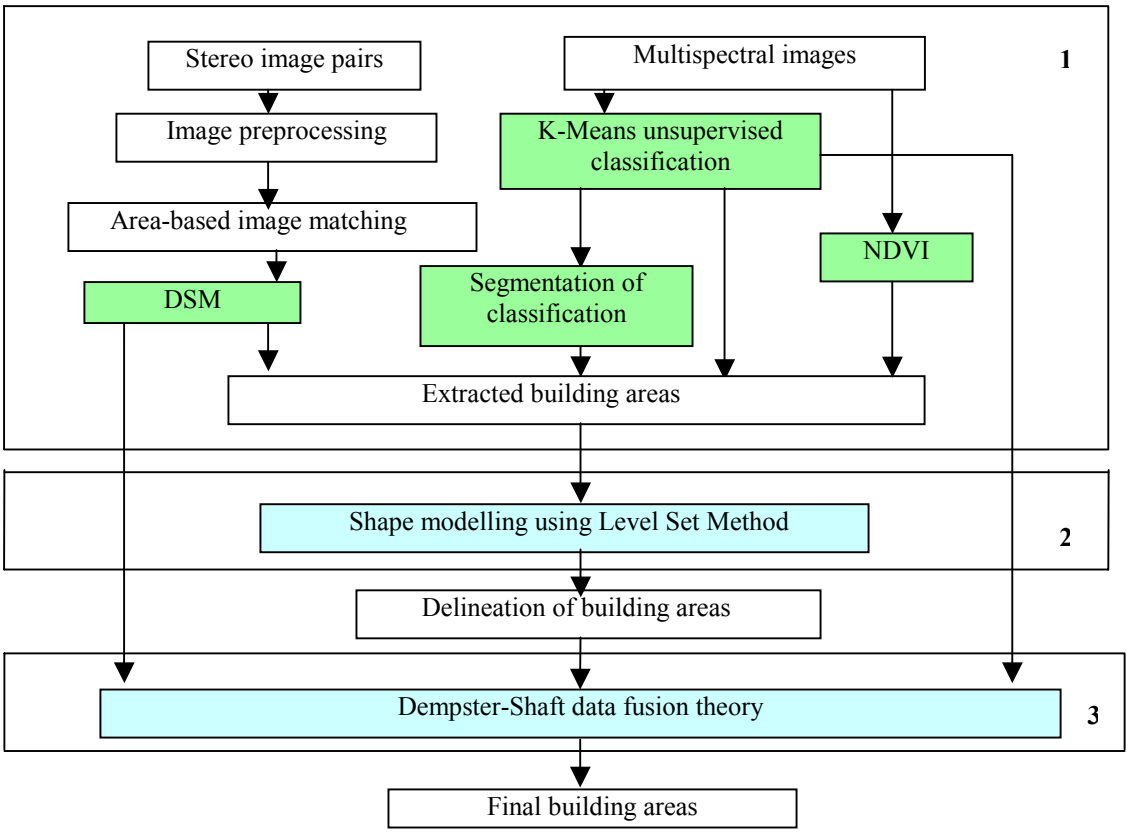

Figure 1 Architecture of the building extraction system

PROCESSING OF STEREO AND MULTISPECTRAL IMAGES

Stereo Image Matching For DSM

Since this paper concentrates on the process of recognizing building in images, the DSM obtained from LH Systems' Socet Set v4.2 has been directly used in the subsequent stages of the system in Figure 1. A dense sample of points in the DSM is obtained 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 Analysis

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.

Extraction Building Interest Areas

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.

SHAPE MODELLING AND IMAGE SEGMENTATION WITH LEVEL SET METHOD

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.

Consider a closed curve moving in a plane. Let $\gamma(0)$ be a smooth, closed initial curve in Euclidean plane R^2, and let $\gamma(t)$ be the one-parameter family of curves generated by moving $\gamma(0)$ along its normal vector field with speed $F(K)$, K is a given scalar function of the curvature. Let $\boldsymbol{x}(s,t)$ be the position vector which parameterises $\gamma(t)$ by s, $0 \le s \le S$.

The level set method represents the front $\gamma(t)$ enclosing curve, as the level set $\{\Phi = 0\}$ of a function Φ. Thus, given a moving closed hypersurface $\gamma(0)$, we wish to produce a formulation for the motion of the hypersurface propagating along its normal direction with speed F. F can be a function of various arguments, including the curvature, normal direction, etc. The main idea is to embed this propagating interface as the zero level set of a higher dimensional function Φ. Let $\Phi(\boldsymbol{x}, t=0)$, where $\boldsymbol{x} \in R^N$ in N dimensional space, is defined by

$$\Phi(\boldsymbol{x}, t = t_1) = \pm d \qquad (1)$$

where d is the distance from $\boldsymbol{x}$ to $\gamma(t=0)$, and the plus sign is chosen if the point $\boldsymbol{x}$ is outside the initial hyersurface $\gamma(t=0)$, the minus sign is chosen if the point $\boldsymbol{x}$ is inside the initial hypersurface. Thus, the initial function $\Phi(\boldsymbol{x}, t=0) : \boldsymbol{x} \in R^N$ can be defined as follows:

$$\gamma(t=0) = (\boldsymbol{x} \mid \Phi(\boldsymbol{x}, t=0) = 0) \qquad (2)$$

Now, we need to produce an equation for the evolving function $\Phi(\boldsymbol{x}, t)$ which contains the embedded motion of $\gamma(t)$ as the level set $\{\Phi = 0\}$.

Using the chain rule in Sethian (1999, 1995), the evolution equation for Φ can be a type of Hamilton-Jacobi equation. Based on the advantages of the Hamilton-Jacobi equation, in two space dimension, a numerical approximation for the evolving function can be obtained. Using the forward and backward difference approximations in Φ, the evolving function can be described as Equation (3) and n defines iterations.

$$\Phi_{ij}^{n+1} = \Phi_{ij}^{n} - F_0 \Delta t((\max(D_x^- \Phi_{ij}, 0))^2 + (\min(D_x^+ \Phi_{ij}, 0))^2 + (\max(D_y^- \Phi_{ij}, 0))^2 + (\min(D_y^+ \Phi_{ij}, 0))^2)^{1/2} - \Delta t F_1 \mid \nabla \Phi \mid \qquad (3)$$

where D_x^+ computes the new values at j using information at j and j+1;

$$D_x^+ \Phi_{ij} = \Phi_{i(j+1)} - \Phi_{ij}$$

similarly D_x^- computes the new values at j using information at j and j-1;

D_y^+ computes the new values at i using information at i and i+1;

D_y^- computes the new values at i using information at i and i-1;

The speed function can be as follows:

$$F = F_0 + F_1(k) = k_I(-1.0 - 0.025K)$$

$$k_I(x, y) = \frac{1}{1 + \mid LOG * I(x, y) \mid} \qquad (4)$$

where K is the curvature of level set. LOG is Laplacian of Gaussian operator.

The intrinsic geometric properties of the front curve may be determined from the level function Φ because F_1 is related to curvature K. The above level set approach can be used in high spatial dimensions.

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. Figure 2 gives an example to show how the level set method works. The example demonstrates that when the evolving curve reaches the boundaries of the building, all the points on the curve stop evolving further and the computation is ended.

Figure 2(a) shows the zero crossings over the building areas, which are obtained by convolving the image with a LOG operator. Figure 2(b) shows the zero crossings overlaid on the original image. Figure 2(c) and 2(d) are

the intermediate curves of the level set modelling. Figure 2(e) shows the final curve from the level set modelling. Only the building boundary shown in blue is extracted because the points on the boundary stop the curve from evolving further. The final boundary of the building overlaid on the original image is shown in Figure 2(f).

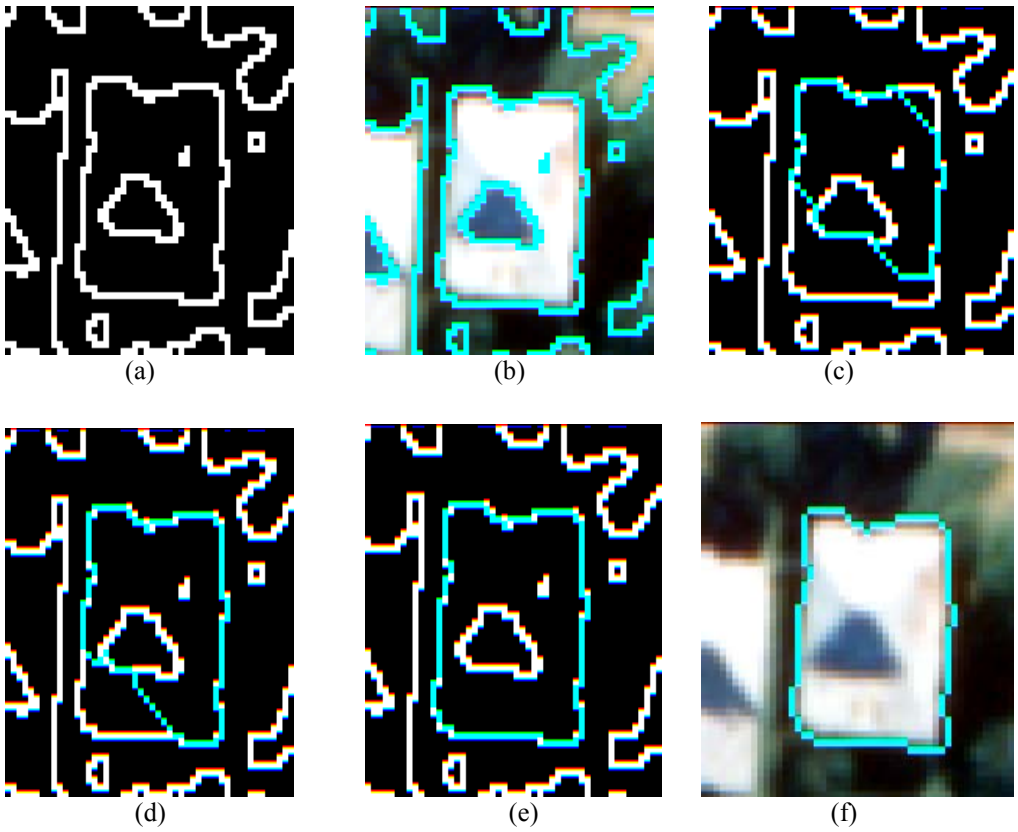

Figure 2 The example of delineation the boundary of the building area

DATA FUSION USING DEMPSTER-SHAFER THEORY

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 **m**(A) is in the interval [0,1].

The following mass equations can be obtained:

$$\mathbf{m}(\varnothing)=0$$
$$\mathbf{m}(\Theta)=\sum_{A\subset 2^{\Theta}}\mathbf{m}(A)=1 \qquad (5)$$

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 $\forall A \in 2^{\Theta}$ and $\forall B \in 2^{\Theta}$, the two parameters are defined respectively as follows:

$$\mathrm{Sup}(A)=\sum_{B\subseteq A}\mathbf{m}(B)$$
$$\mathrm{Pls}(A)=\sum_{B\cap A\neq\varnothing}\mathbf{m}(B) \qquad (6)$$

An uncertainty interval is defined by [Sup(A), Pls(A)] where

$\mathrm{Sup}(A) \leq \mathrm{Pls}(A)$

$\mathrm{Pls}(A)=1-\mathrm{Sup}(\overline{A})$, $\quad A \cup \overline{A} = \Theta$, $A \cap \overline{A} = \varnothing$ (7)

The support value of hypothesis A may be interpreted as the minimum uncertainty value about A. Its plausibility may be interpreted as the maximum uncertainty value of A. The uncertainty interval gives a measurement of the imprecision about the uncertainty value. For several data sources, the Dempster-Shafer method allows compatible propositions to combine the probability masses from these sources to obtain a single value for the probability of the intersection (union) of the propositions.

Assume there are two data sources, A_i object types in data source 1, and B_j object types in data source 2. The object type A is the subset of A_i or B_j. This means A is one type of A_i or B_j. The total probability mass committed to a subset A from two data sources is

$$\mathbf{m}(A)=\frac{\sum_{\substack{i,j \\ A_i \cap B_j = A}} \mathbf{m}_1(A_i)\,\mathbf{m}_2(B_j)}{1-K}, \quad K \neq 1$$

where
$$K = \sum_{\substack{i,j \\ A_i \cap B_j = \varnothing}} \mathbf{m}_1(A_i)\,\mathbf{m}_2(B_j) \qquad (8)$$

For more data sources, $\mathbf{m}_t$ is the basic probability mass provided by source t ($1 \leq t \leq p, p \geq 3$), the combination of all the data sources is defined as follows:

$$\mathbf{m}(A)=\frac{\sum_{A_1 \cap A_2 \ldots\ldots \cap A_p = A} \prod_{1 \leq t \leq p} \mathbf{m}_t(A_i)}{1-K}, \quad K \neq 1$$

where
$$K = \sum_{A_1 \cap A_2 \ldots\ldots \cap A_p = \varnothing} \prod_{1 \leq t \leq p} \mathbf{m}_t(A_i) \qquad (9)$$

In the Dempster-Shafter theory, the hypotheses about single classes and hypotheses about unions of classes are respectively called simple hypotheses and compound hypotheses. When the probability mass of simple hypotheses are not null, a decision rule must be determined that best suits the application, such as the maximum support over simple hypotheses. The formula is as follows:

$$\begin{aligned} &max(Sup(A)) \\ &Sup(A) \geq Sup(\overline{A}) \end{aligned} \qquad (10)$$

This method has been used to combine the three datasets of building classification, level set result and DSM, to extract the final reliable building areas. The evaluation procedure by Dempster-Shafer evidential reasoning described in this study is based on spatial features. This means that the determination of probability masses and their combination is based on features. For each region derived by LevelSet, there are corresponding areas obtained from the DSM and clustered image. The clustered image, regions of LevelSet and DSM are assigned as data sets 1,2, and 3 respectively. Classes C1, C2 and C3 represent trees, buildings and ground respectively.

Table 1 shows the computed probability mass, *plausibility* and *support* values for each simple and compound hypothesis based on three data sources. For

Table 1: Dempster-Shafer calculation from three data sources

$$1-k = 4(ut + st + us - 9uts)$$

A	$\boldsymbol{m}_1$(A) clustered image	$\boldsymbol{m}_2$(A) regions LevelSet	$\boldsymbol{m}_3$(A) DSM	$\boldsymbol{m}$ (A)	Sup(A)	Sup($\overline{\mathrm{A}}$)
C1 trees	$1-3t$	u	s	$\frac{4us(1-3t)}{1-k}$	$\frac{4us(1-3t)}{1-k}$	$\frac{4t(s+u-6su)}{1-k}$
C2 buildings	t	$1-3u$	s	$\frac{4st(1-3u)}{1-k}$	$\frac{4st(1-3u)}{1-k}$	$\frac{4u(s+t-6st)}{1-k}$
C3 ground	t	u	$1-3s$	$\frac{4ut(1-3s)}{1-k}$	$\frac{4ut(1-3s)}{1-k}$	$\frac{4s(u+t-6ut)}{1-k}$
C1 $\cup$ C2	0	0	s	0	$\frac{4s(u+t-6ut)}{1-k}$	$\frac{4ut(1-3s)}{1-k}$
C1 $\cup$ C3	0	u	0	0	$\frac{4u(s+t-6st)}{1-k}$	$\frac{4st(1-3u)}{1-k}$
C2 $\cup$ C3	t	0	0	0	$\frac{4t(s+u-6su)}{1-k}$	$\frac{4us(1-3t)}{1-k}$
Θ	0	0	0	0	1	0

data set 1, clustered image, the probabilities for all the simple and compound classes are represented by probability *t,* as an example. Similarly, for regions derived from LevelSet and DSM, the probabilities for the classes are represented by *u* and *s* respectively.

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. For data set 1, the clustered image, $m_1(A)$ there are no null values assigned to C1, C2, C3 and $C2 \cup C3$ as shown in Table 1. Since C2 for buildings and C3 for the ground may have the same texture in the image, there are ambiguities between these two classes in clustered image. In Table 1, C2, C3 and $C2 \cup C3$ are assigned the same probability, *t*. Null probability masses are assigned to the other compound hypotheses. Since all the masses sum to 1, the probability mass for C1 is $1 - 3t$.

For each extracted region of LevelSet, the numbers of pixels representing as trees, ground, glasses and building can be calculated respectively. The probability *u* can be defined based on pixel assigned to each building region. For data set 2, regions derived by LevelSet, since the extracted building regions are derived from the processing of low-level image analysis and interpretation and level set modelling based image segmentation, they are more reliable and are assigned higher probabilities. As shown in Table 1, buildings in class C2 are assigned a probability of $1 - 3u$ and other non-null classes are assigned a probability of *u*. If the other classes are assigned lower probabilities, the class of building will have higher probability since the sum of probabilities is 1.

For data set 3, DSM image, since buildings and trees are above ground areas in the DSM, when they have about the same height, these two classes areas are easily confused. As shown in the Table 1, class C1 representing trees, C2 of buildings and $C1 \cup C2$ are therefore assigned the same probabilities, *s*. The probability mass of class C3, the ground is *1-3s*. The calculation of *s* is based on the mean of the DSM values. For example, if the mean of elevations for a building area is high, the probability *s* assigned to that building area in the DSM is assigned as higher than other areas.

Table 2 is example to show how one building can be evaluated based on three datasets. Based on the decision rule Equation (10), since the decision is determined by simple hypotheses, the processed building has been evaluated as buildings from three data sets as shown in the values of the last two columns.

Table 2 One building evaluation using three datasets: clustered image, regions of LevelSet and DSM

t=0.16, u=0.15, s=0.33, 1-k=0.22

A	m_1(A) clustered img	m_2(A) regions LS	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	0.52	0.15	0.33	0.47	0.47	0.53
C2 buildings	0.16	0.55	0.34	0.53	0.53	0.47
C3 ground	0.16	0.15	0	0.01	0	1
$C1 \cup C2$	0	0	0.33	0	1	0
$C1 \cup C3$	0	0.15	0	0	0.47	0.53
$C2 \cup C3$	0.16	0	0	0	0.53	0.47
Θ	0	0	0	0	1	0

TESTS AND RESULTS

While a number of test areas have been investigated, the results of processing one area will be given in the following. Figure 3 illustrates a left image of a pair of colour aerial images with 522 × 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. Figure 4 illustrates the DSM map obtained from the stereo image matching using LH Systems' Socet Set. Figure 5 is the result of unsupervised classification using the K-Means method. Yellow is assigned to building areas, red and green present the vegetation areas and blue is considered to be road areas. Based on the classification of Figure 5, using a post classification procedure, a segmented image can be created to show the areas classified as buildings. Most of the building areas have been detected, but the dark

Figure 3 Left image of stereo aerial images

Figure 4 DSM from stereo image matching

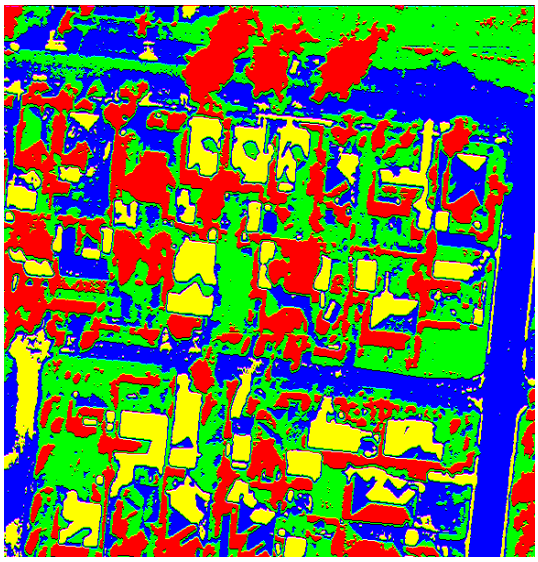

Figure 5 Unsupervised classification by clustering

Figure 6 Building interest areas after image analysis

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 four information layers of DSM, classification, post classification segmentation and VVI are input into the ArcView, using the Map Queries operation, and the possible building areas are extracted. Using the region growing algorithm, small spots which do not belong to buildings can be deleted from the Map Queries result. The building interest areas are derived by analyzing the multispectral and stereo images. The building interest areas are overlaid on the ortho image as shown in Figure 6. Some road areas wrongly assigned as buildings have been deleted and the correct building areas derived from classification have been successfully maintained.

From Figure 6, the approximate building areas can be obtained. For every small building area, shape modelling with level set method was then implemented. The delineated boundaries of the buildings are shown Figure 7. The level set results overlaid on the ortho image are shown in Figure 8. Some regions in the left of the image, which belong to road areas, are still assigned as building areas after the level set shape modelling. This is because the building interest areas supply wrong information and cause some building regions to be unreliable. Thus, it is necessary to use Dempster-Shafer data fusion method to evaluate the regions.

The Dempster-Shafer method provides a single value for the probability of the intersection (union) of the propositions. Based on the decision rule given in Equation 10, the three data sources, clustered image, regions of LevelSet and DSM were combined to produce reliable building areas shown in Figure 9. The final building boundaries overlaid on ortho is shown in Figure10.

The consequence of the data fusion is that three incorrect building areas have been detected and deleted in the final result, as shown in Figure 9. The correct building areas shown in Figure 7 are unchanged after data fusion in Figure 9. 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.

CONCLUSIONS

The method described in this paper combines stereo image matching, multispectral image analysis, shape modelling by the level set method and Dempster-Shafer data fusion theory, to locate building areas in the test images. The level set technique can be used to aid the tasks of object representation and recognition, but it needs the user to provide an initial contour. In order to automate the extraction of buildings, the DSM and multispectral image analysis can supply approximate building areas. These building interest areas can be used to determine the initial contours for the level set method and overcome the requirements for user input. This method has been applied for the first time in this study for complicated urban image analysis. The multiple data sources result in the modelling of shapes for extracted building interest areas. This method overcomes the

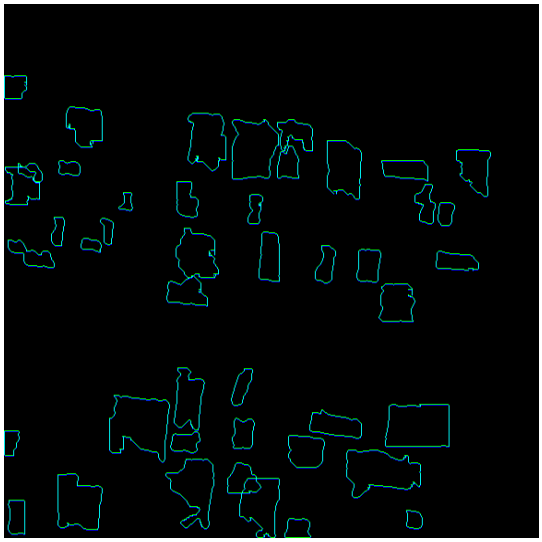

Figure 7 The results of building boundary by level set

Figure 8 Fig. 7 overlaid on ortho image

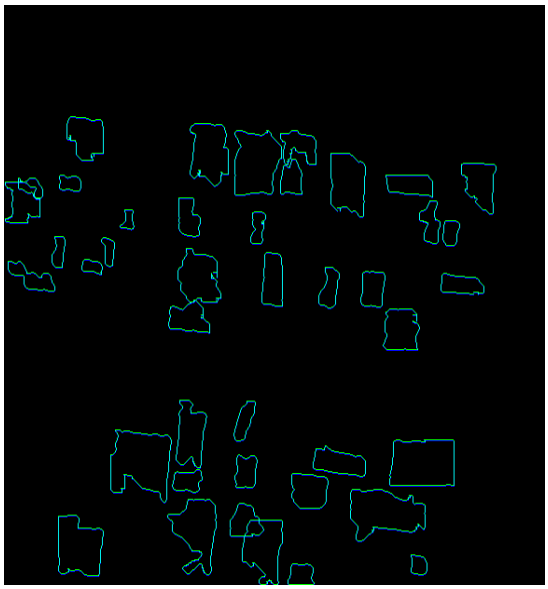

Figure 9 Result of Dempster-Shafer based on 3 datasets

Figure 10 Fig. 9 overlaid on ortho image

complicated and difficult 2D inference in a single image. The Dempster-Shafer data fusion technique provides the theoretical basis for evaluating the reliability of the extracted buildings from the combination of the different data sources by a statistically-based classification. Based on the one test area shown, the results are encouraging, but further research is needed to refine these methods of extract building interest areas, in order to extract all the buildings with different textured roofs.

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. 1995. 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.