FAST INFORMATION ACQUISITION ON BUILDING FEATURES FROM DIGITAL AERIAL IMAGES

M. Caprioli*, B. Figorito*, E. Tarantino*

DVT- Polytechnic University of Bari -Italy
E-mail: m.caprioli@poliba.it; benedetto1980@libero.it; e.tarantino@poliba.it

ABSTRACT

In the domain of emergency response to natural disasters and man-made conflict events, rapid situation assessment is crucial for initiating effective emergency response action. Estimation of damages caused by a disaster is a major task in the post disaster mitigation process.

To enhance the relief and rescue operation in the affected area it is required to get a near real time damage model. For this purpose a fast method of data acquisition with suitable methods for extracting the man-made objects is required.

In this work a methodological protocol is proposed to generate building description which can be used in a simulation system for training emergency forces. In this regard it is sufficient to extract 2D roof outlines of the buildings and attribute them with a constant height value, resulting in flat roof buildings. The process of building extraction from imagery can be separated in two tasks: first the detection of the object, second the reconstruction. Another focus of this research is on the development of methodologies to extract automatic nDSM in order to remove non-ground objects such as trees, and cars from the dataset.

KEYWORDS: Building extraction, stereo images, digital surface model, data fusion, segmentation, geo-processing

1. INTRODUCTION

The objective of disaster management is to mitigate the effects of disasters by effective use of available resources, i.e. for example to organize the activities of rescue teams (Markus et. al., 2000) or to provide an information base for the responsible executives. Probably the most important prerequisite for these strategies is a concept for fast information acquisition and frequent updates, since the situation after a disaster must be known to initiate appropriate counter-measures.

The idea for damage interpretation is to detect changes at buildings by comparing pre- and post-disaster building models (Steinle et. al., 2001). If there are no 3D city models available, the pre-disaster building models can be extracted and modelled from digital aerial imagery.

The automation in features extraction (i.e. buildings, roads, trees, etc.) is one of the main prolific field of scientific research. In literature there are many approaches and strategies that makes difficult an effective comparison. Problems are still unsolved and depend prevailingly on spatial resolution of input data. The main goals of this research field are the reduction of time and costs in production and updating of topographic database. A particular and advanced automation example regards the feature extraction software tool allowing the cartographic entities extraction replacing the human ability in digitalizing, interpretation and evaluation.

The same procedure if applied at images with a different Ground Sample Distance (GSD) could carry out at totally different results. The term automation is always used in a deceitful manner and such approximation does not allow the execution of a true comparison among different methodologies.

As Gerke et al. (2001) stated, the process of building extraction from imagery can be separated in two main tasks: first the detection of the object ("Where is a building?"), second the reconstruction ("Which geometrical description can be found for this object?"). Regarding these topics a lot of publications can be found. Some of these deal with the complete process, others concentrate on one of the named tasks.

In this work we refer to the automatic classification systems introduced by Guelch (2000). For such systems the main phases are conducted in automatic manner, the user is involved in phases related to initial setup, post-editing and in errors identification in final results.

For this reason in cartographic applications and firms semi-automatic systems are prevailing because it is difficult distribute with many errors data.

The consolidated methods vary for building typology, detail level, images number, geometric primitives used, a priori information and automation/ user interface level. Data and sensors are generally at high resolution (GSD < 1 m) allowing single building identification. The height information is obtained using stereo images or a Digital Surface Model (DSM).

2. DATA AND METHODS

The image data used in this work have a 22 cm GSD as distance on the ground, represented by each pixel in the x and y components, and were acquired on 18 August 2008 using the digital photogrammetric camera Z/I Imaging DMC – Intergraph.

To reduce processing times and because of the peculiar location, such data represent a sub-set area overlapping two image data (sizes 2244 x 1473 pixels – 13 ha of surface). The study area is located at the Marconi - San Girolamo-Fesca quarter on the North-West of Bari (Italy), amid the Lama Balice and the Bari Palese airport (Fig. 1). The vegetation close to the many three-floor residential buildings and the multiform roofs typology of various materials were some of the difficulties to overcome which motivated this work.

The proposed method has a high flexibility with a lonely restriction connected to the use of stereo images for the third dimension reconstruction. Such constrain depends on the precise choice of not using digital terrain models derived from others raw data. The present approach was conceived for images acquired with nadiral view from aerial platform. It could

be used also with an oblique acquisition view or with satellite images by making further adaptations.
The methodology is structured in five definite phases: pre-processing operations; DSM and orthophoto generation; image processing; nDSM generation; buildings identification and extraction.

Fig. 1. The study area

2.1 Pre-processing operations

This phase was prevailingly oriented towards the stereoscopic model creation by using the Leica Photogrammetry Suite 9.2 (LPS) software tools.

2.2 DSM and orthophoto generation

The Digital Surface Model was created with the LPS Automatic Terrain Extraction (ATE) add-on module, by setting up 0.85 as limit value of the correlation coefficient and the research strategy "high urban" for the determination of the homologous points in the images. This procedure is fully automatic, without manual digitations, breaklines and post-editing intervention, establishing a 2 m regular grid and points with optimal correlation. After the determination of the altimetric model, the orthophoto was generated imposing 22 cm as grid cell dimension with a bilinear resampling.

2.3 Image processing

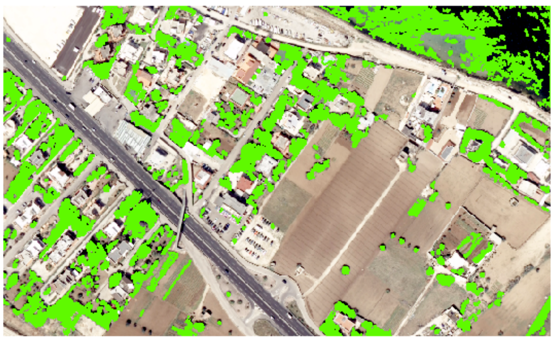
To improve the information extraction in this phase the resulted orthophoto was treated with three distinct transformations associated with their respective layers (*Shadow, Vegetation* and *Canny-Edge*).

a) The Shadow Layer
The creation of this layer facilitated the identification of the shadow zones, determining no-building areas. To such purpose, the resulted orthophoto was transformed from RGB to HSV (Hue, Saturation e Value), as many authors suggested (Salvador et al., 2001; Tsai, 2006). Small shadow zones, with staircases, trees, electric wires etc. were considered as illuminated ones.
The Shadow Layer was obtained executing the product of two masks: the first mask was derived from the saturation component and the second one was extracted on the intensity component, with 0.5 as minimum value for both of them. Moreover, all the shadow zones with a smaller surface of 8 sq. m. (about 200 pixels) were removed. This mask had many large false positives, that is no-shadow zones identified as shadow ones. In this method features with low digital numbers in the Blue band (deep water bodies, some asphalt typologies, cars, vegetation) were considered important for building extraction (Fig. 2).

Fig. 2. The Shadow Layer represented with blue colour

b) The Vegetation Layer
With the Vegetation Layer it was possible to distinguish plant biomass using the spectral contents of the RGB data, because of the absence of the Infrared information. In this way the Red and Green bands were combined according to the method suggested by
Sibiryakov (1996). The Infrared band was simply substituted with the Green band in order to constitute the NDGRI (Normalized Difference Green Red Index). To create the NDGRI mask a minimum threshold value of 0.2 was imposed. Such value was empirically determined by means of the Spectral Signatures analysis on single ground entities.
The main limit of the NDGRI index is related to erroneous classification of some roof typologies. To reduce these effects the texture homogeneity parameter of the GLCM (Grey Level Co-occurrence Matrix) was introduced. For the homogeneity mask a maximum threshold value of 0.5 was imposed on the base of the empirical analysis related to the plant biomass in the study area.
The spectral information of both the NDGRI and homogeneity mask layers were then combined making an intersection processing, by selecting only the feature area with surfaces greater than 4 sq. m. (about 100 pixels).
Lastly, the dilation morphological operator with a square kernel of 5 pixels was applied to reduce the presence of holes in polygons and include the vertexes number of every feature area (Fig. 3).

Fig. 3. The Vegetation layer represented with green colour



c) The Canny Edge Layer
The purpose of the Canny Edge Layer creation was the edge detection, which aims at identifying points in a digital image at which the image brightness changes sharply or more formally has discontinuities. Such discontinuities are likely to correspond to: discontinuities in depth, discontinuities in surface

orientation, changes in material properties and variations in scene illumination.
In literature there are many edge detection algorithms (Ziou, et al. 1998), with the aim of reducing the number of false reconnaissance produced by noises in images.
In this work the Canny algorithm (1998) was used because of its efficaciousness in edge detection. The Canny algorithm contains a number of adjustable parameters, which can affect the computation time and effectiveness of the algorithm.
The first parameter is the size of the Gaussian filter. The smoothing filter used in the first stage directly affects the results of the Canny algorithm. Smaller filters cause less blurring, and allow detection of small, sharp lines. A larger filter causes more blurring, smearing out the value of a given pixel over a larger area of the image. Larger blurring radii are more useful for detecting larger, smoother edges - for instance, the edge of a rainbow.
The second parameter is Thresholds. The use of two thresholds with hysteresis allows more flexibility than in a single-threshold approach, but general problems of thresholding approaches still apply. A too high threshold set can lack in important information. On the other hand, a too low threshold set will falsely identify irrelevant information (such as noise) as important. It is difficult to give a generic threshold that works well on all the images. No tried and tested approach to this problem yet exists.
The Canny algorithm was applied on the band Value of the HSV colour space, imposing 1.5 as the value of the standard deviation in the Gauss smoothing filter and 3 as lower threshold value for the non-maximum suppression (Fig 4).

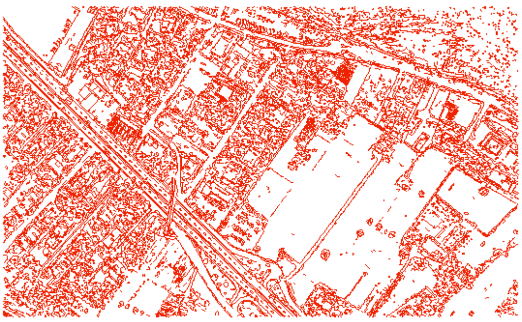


Fig. 4. The Canny-Edge Layer represented with red colour

2.4 nDSM generation

nDSM is a type of raster GIS layer which represents a regular arrangement of locations and each cell has a value corresponding to its elevation, that is normalized or “untied” from the ground.As the nDSM represents the absolute elevation of artificial elevation features regard to the earth, for preparing of nDSM the DSM should be subtracted from DEM, so that points laid on the ground have values 0, while trees and buildings have values greater than 0.
The nDSM procedure was divided in three phases.
The first phase was based on the determination of 3D points with a searching strategy different from “high urban”/DSM generation and comparing the results of both the strategies (Gooch et al., 2001). The second cloud points was generated using the “low urban”. In this way if a point have the same height value, including the Root Mean Squared Error, with a high probability it is a ground point. Then a segmentation of the difference between the two altimetric models was executed, imposing 0.05 as edge threshold d, a city-block metric for the contours and a number of pixels for every feature area not lower than 200.
In the second phase of the procedure the filtering to obtain points laid on the ground was applied, selecting by means of spatial queries only the points included in the feature areas previously extracted.
Lastly, the difference between points resulted from “high urban” strategy and points laid on the ground was executed to obtain the nDSM model (Fig. 5).

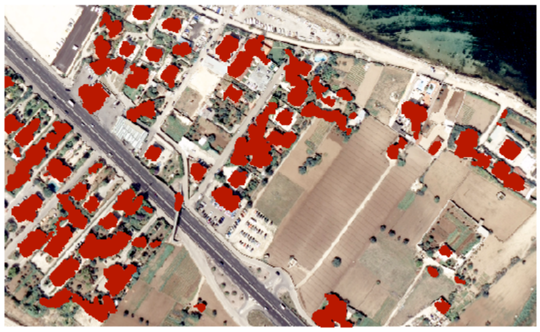


Fig. 5. The nDSM model

Then other two layers based on the nDSM model were created.
The Elevation Layer (Fig. 6) was extracted by means of a segmentation procedure with 2.5 m as threshold edge, a chessboard metric for contours processing, a pixels number greater than 900 and using the *Opening* morphological operator with kernel 3x3.

Fig. 6. The Elevation Layer

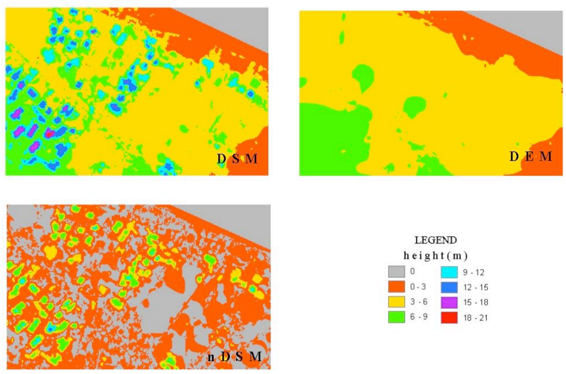


Fig. 7. The Slope Layer

The Slope Layer (Fig. 7) was processed using the slope data and imposing a segmentation with 40° as threshold edge, a chessboard metric for contours processing, a pixels number greater than 180 and using the *Erode* morphological operator with kernel 3x3.

2.5 Buildings identification and extraction

The Buildings identification and extraction procedure was the most sensitive to the alterations of the introduced parameters connected to the local conditions of the acquired scene. In this last phase of the proposed method all the previously derived information were implemented in a GIS environment using the Model Builder tool of the ArcGis 9.2 software (Fig. 8). The parameters definition depends on the user a priori knowledge (e. g. size and mean height of buildings) and could include the modification of the minimum area, the thickness, etc.

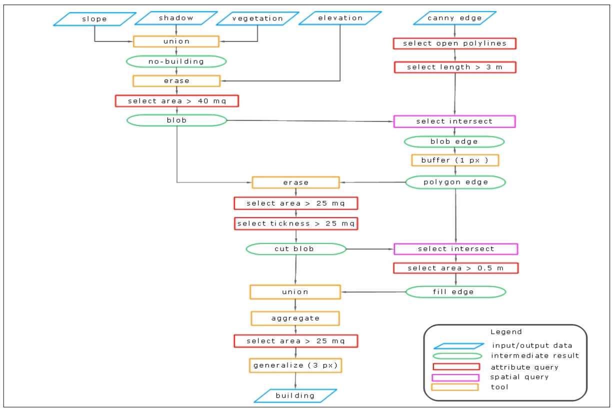


Fig. 8. The workflow of the proposed method

The Shadow, Vegetation, Elevation and Slope layers were therefore combined to determine the no-building areas. The Canny-Edge Layer was used to finish polygons separating the areas close to buildings from the internal ones (Fig. 9).

Fig. 9. The blue checks correspond to no-building areas while the red lines are the Canny-Edge features

In the intermediate operations spatial and attribute queries were executed to filter only the polygons potentially able to identify buildings. With the *Union* and *Aggregate* procedures the internal features with the buildings contours were fused removing the internal holes. The *Generalize* procedure (with a tolerance of 3 pixels) allowed the reduction of the number of the vertexes polygon. Lastly, the buildings were extruded by considering the height mean value of the DSM included in every building (Fig. 10).

Fig. 10. 3D view with the extruded buildings shape on the processed orthophoto

3. RESULTS AND CONCLUSIONS

The proposed buildings extraction methodology was qualitatively verified draping the Digital Building Model (DGM) on the orthophoto (Fig. 11).

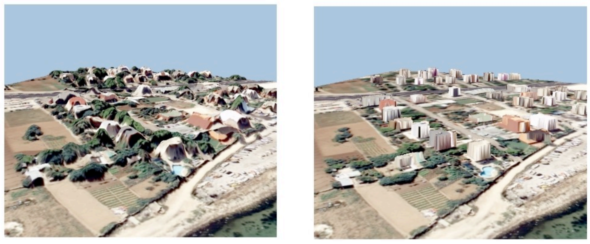


Fig. 11. Orthophoto draped on the DSM (left) and DEM with building extruded with orthophoto (right)

From the quantitative point of view, the results were compared using the following Detection Percentage (DP) and Branch Factor (BF) indexes (Lin and Nevatia, 1998):

(1)

(2)

The DP index describes in percentage the number of buildings correctly recognised by means of the automatic procedure while the BF index expresses the number of buildings erroneously identified. (Table 1)

TP	49
FP	1
TN	3
DP	94.2 %
BF	2.0%

Table 1. Values of the indicators and indexes for the quantitative assessment of the results

Such indexes were calculated by means of the comparison with ground true data obtained with the photogrammetric plotting. In this verification phase various indicators were applied to obtain the DP and BF indexes. The True Positive (TP) represented a correctly identified building by the user and with the automatic procedure. The False Positive (FP) was a building recognised with the automatic approach but not by the user. The True Negative (TN) indicator was a building identified by the user but not with the automatic procedure.

The Fig. 12 shows the results obtained considering 53 buildings in the test area.

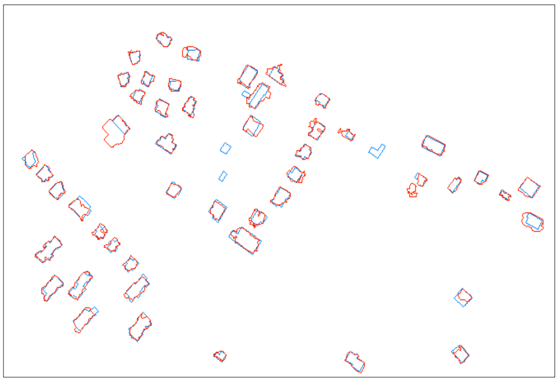


Fig. 12. Results of the quantitative comparison

The proposed automatic building extraction method from digital aerial images showed good results if we compare them with similar works in this application field. The second and fourth phases generated some matching errors checkable in the final results. They were due to presence of occlusions, vegetation, similar spectral content among covers and soil.
The different building typologies in this study area proved that in areas with major building regularity the proposed methodology could guarantee better results.
Further studies are foreseen with other source data (e.g. satellite) and territorial features to better validate the proposed method.

REFERENCES

Baltsavias E.P., Gruen A., Gool, 2001. Automatic Extraction of Man-Made Objects fron Aerial and Space Images (III). A.A. Balkema Publishers.

Canny, J., 1986. A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8, pp. 679-714.

Gerke, M., C. Heipke, and B. M. Straub, 2001. Building extraction from aerial imagery using a generic scene model and invariant geometric moments, *Proceedings, IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, 08–09 November, Rome, Italy, pp. 85–89.

Gooch, M. J., Chandler J. H., 2001. Failure prediction in automatically generated digital elevation models. Computers & Geosciences 27(8), pp. 913-920.

Guelch E., 2000. Digital Systems for Automated Cartographic Feature Extraction. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B2, pp. 241-256.

Lin C., Nevatia, R., 1998. Building detection and description from a single intensity image". Computer Vision An image Understanding CVIU 72(2), pp. 101-121.

Markus, M., Fiedrich, F., Gehbauer, F. & Hirschberger, S. 2000. Strong Earthquakes, Rapid Damage Assessment and Rescue Planning. In Kowalski, K.M. & Trevits, M.A. (eds.), *Contingency, Emergency, Crisis, and Disaster Management: Emergency Management in the Third Millennium; Proc. 7th Annual Conference of the International Emergency Management Society, 369-378, Orlando, May 1999.*

Sibiryakov A., 1996. House detection from aerial color images. Internal Report, Institute of Geodesy and Photogrammetry, ETH Zurich.

Salvador E., Cavallaro A., Ebrahimi, T., 2001. Shadow identification and classification using invariant colour models. IEEE International Conference on Acoustic, Speech, and Signal Processing, pp. 1545-1548.

Steinle, E. & Voegtle, T., 2001. Automated extraction and reconstruction of buildings in laserscanning data for disaster management. In: *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), E. Baltsavias et al. (eds.),* Swets & Zeitlinger, Lisse, The Netherlands, pp. 309-318.

Tsai V.J.D., 2006. A comparative study on shadow compensation of color aerial images in invariant color models. IEEE Transactions on Geoscience and Remote Sensing 44(6), pp. 1661-1671.

Weidner U.,Foerstner W., 1995. Towards Automatic Building Extraction from High-Resolution Digital Elevation Models. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 50, Issue 4, pp. 38-49.

Ziou, D. and Tabbone, S., 1998. Edge Detection Techniques An Overview. International Journal of Pattern Recognition and Image Analysis, 8(4) pp. 537-559