

Estrazione automatica di edifici da immagini aeree

Mauro Caprioli, Benedetto Figorito, Eufemia Tarantino,

Politecnico di Bari, Via Orabona 4 70125 Bari, tel. 080 5963732, fax 080 5963719
e-mail: m.caprioli@poliba.it, benedetto1980@libero.it, e.tarantino@poliba.it

Riassunto

Il principale obiettivo della presente ricerca è sviluppare una procedura automatica per l'estrazione di edifici da immagini aeree ad alta risoluzione. La metodologia proposta ha un'alta flessibilità e una sola restrizione che è legata all'utilizzo di immagini stereo per la ricostruzione della terza dimensione. Questo vincolo dipende dalla precisa scelta di non utilizzare modelli digitali del terreno calcolati a partire da altri fonti di dati. L'intera procedura è strutturata in cinque parti: operazioni preliminari, creazione DSM ed ortofoto, elaborazioni immagini, creazione nDSM, individuazione ed estrazione degli edifici. I risultati conseguiti sono confrontabili con quelli ottenuti con altri metodi presenti in letteratura, ma ancora non direttamente utilizzabili per la produzione cartografica se non attraverso processi di *post editing*.

Abstract

In this work an automatic procedure for buildings extraction was mainly implemented from high resolution airborne imagery in order to evaluate its reliability. The proposed methodology has high flexibility and the constrain connected to the use of stereo images for the third dimension reconstruction. Such constraint depend on the choice of not using digital terrain models based on other raw data. The methodology is structured in five phases: pre-processing, DSM and orthoimage generation, nDSM creation, buildings identification and extraction. The obtained results with this method are comparable with others made with other previous tested methods, but still not directly usable for cartographic production except with post-editing processing.

Introduzione

L'automazione delle procedure è un'area di ricerca in continua evoluzione, trainata dal desiderio continuo di ridurre tempi e costi nella produzione e nell'aggiornamento dei *database* topografici. Negli ultimi anni infatti, e in particolare con il lancio dei primi sensori ad alta risoluzione, sono stati richiesti, ai ricercatori e alle aziende produttrici di dati cartografici, degli enormi investimenti nell'automazione. Un esempio di automazione tra i più avanzati riguarda lo strumento di *feature extraction* disponibile in alcuni software che consentono di estrarre le entità cartografiche sostituendosi all'uomo nel fornire digitalizzazione, interpretazione e valutazione.

L'automazione nella estrazione di entità come edifici, strade e singoli alberi, è uno dei filoni più prolifici nella ricerca scientifica, i problemi però non appaiono del tutto risolti. Il termine automazione è spesso utilizzato in maniera ingannevole, tale approssimazione non consente di effettuare un vero confronto fra le diverse metodologie. In letteratura esistono vari approcci e varie strategie che rendono difficoltoso un reale confronto fra le diverse metodologie proposte: le varie procedure dipendono prevalentemente dalla risoluzione spaziale dei dati di *input*. La stessa procedura applicata ad immagini aventi un GSD (*Ground Sample Distance*) diverso può portare a risultati completamente diversi. La classificazione del grado di automazione è quindi fondamentale per confrontare tra loro le diverse procedure.

I ricercatori di tutto il mondo concentrano i loro studi nei sistemi automatici, però non sono in grado di garantire la qualità dei prodotti finali. Per questo motivo nelle aziende cartografiche e nelle

diverse applicazioni, tali sistemi cedono il passo ai sistemi semi-automatici perché è difficile immaginare di poter distribuire risultati finali con una percentuale non trascurabile di errore.

Le metodologie presenti in letteratura per l'estrazione automatica degli edifici variano in funzione della tipologia di edificio, del livello di dettaglio, del numero delle immagini, delle primitive geometriche utilizzate, delle informazioni a priori, del livello di automazione/interfaccia operatore.

I dati utilizzati nelle diverse metodologie per l'estrazione di edifici sono generalmente ad alta risoluzione ($GSD < 1$ m), poiché tale proprietà geometrica consente di individuare singoli edifici. L'informazione relativa alla quota viene fornita utilizzando delle immagini stereo, ovvero utilizzando un DSM (*Digital Surface Model*).

Nel presente lavoro di ricerca si fa riferimento ai sistemi automatici previsti nella classificazione introdotta da Guelch (2000) sui livelli di automazione. Per questi sistemi le fasi principali sono svolte in modo automatico, nelle quali l'operatore cura le fasi relative al *setup* iniziale, al *post-editing* e in generale alla correzione di possibili errori nei risultati finali.

Dati e metodi

L'area test è localizzata a nord-ovest della città di Bari nel quartiere Marconi - San Girolamo-Fesca, tra la foce della lama Balice e l'aeroporto cittadino di Bari Palese. Considerando la dimensione del pixel e la distanza media di presa le immagini risultano con un GSD di circa 22 cm. Tali immagini sono state acquisite il 18 agosto del 2008 utilizzando la camera fotogrammetrica digitale Z/I *Imaging DMC* della *Intergraph*.

I risultati finali della procedura di triangolazione, con il calcolo dei sei parametri per ogni fotogramma, ossia le tre coordinate del centro di presa (X_0 , Y_0 e Z_0) e i tre angoli di rotazioni (ω , ϕ e κ), rispetto al sistema di riferimento Roma40, sono evidenziati in Figura 1.

file	X_0 [m]	Y_0 [m]	Z_0 [m]	ω [gon]	ϕ [gon]	κ [gon]
1288_00762	4556160.915	2671299.659	1996.552	0.16746	399.92329	381.87714
1288_00763	4555970.454	2671820.120	1981.348	0.22379	0.39709	381.95847

Figura 1 – Coordinate dei centri di presa

L'area di studio rappresenta una porzione della superficie di sovrapposizione delle due immagini. Le sue dimensioni sono di 2244x1473 pixel, pari a circa 13 ettari (Fig.2).



Figura 2 - Area test

La scelta di limitare l'area di studio è legata sia ai tempi di calcolo dell'intera procedura sia agli aspetti peculiari della zona che presenta delle abitazioni residenziali con al massimo tre piani.

La presenza della vegetazione a ridosso degli edifici e la tipologia dei tetti, variegata sia nelle forme che nei materiali utilizzati, sono alcuni degli ostacoli da superare nel presente lavoro.

La metodologia proposta per l'estrazione ha un'alta flessibilità, ma una sola restrizione che è legata all'utilizzo di immagini stereo per la ricostruzione della terza dimensione.

Questo vincolo dipende dalla precisa scelta di non utilizzare modelli digitali del terreno calcolati a partire da altre fonti di dati. Questo approccio è stato concepito per immagini acquisite con ripresa nadirale da piattaforma aerea, ma con i necessari adattamenti può essere utilizzato anche con riprese oblique o con immagini satellitari. La metodologia è strutturata in cinque parti definite:

- 1) operazioni preliminari;
- 2) creazione DSM ed ortofoto;
- 3) elaborazioni immagini;
- 4) creazione nDSM;
- 5) individuazione ed estrazione degli edifici.

La procedura di *pre-processing* è stata svolta utilizzando il software di fotogrammetria *Leica Photogrammetry Suite 9.2*. La creazione del DSM è stata effettuata utilizzando il modulo ATE di *LPS*, impostando il valore limite del coefficiente di correlazione a 0.85 e la strategia di ricerca "*high urban*" per l'individuazione dei punti omologhi delle immagini.

L'intera procedura, dato il suo carattere automatico, non prevede la digitalizzazione manuale, l'utilizzo delle *breaklines* e le procedure di *post-editing*. Inoltre è stato impostato un passo della griglia pari a 2 metri e sono stati selezionati i punti aventi uno status di correlazione "*excellent*" o "*good*". L'ortofoto è stata quindi generata avendo determinato in precedenza il modello altimetrico ed impostando la dimensione della cella pari a 22 cm e un ricampionamento bilineare delle stesse. Nella terza fase l'ortofoto è stata trasformata attraverso due operazioni distinte, che hanno consentito, nelle successive parti della procedura proposta, di migliorare l'estrazione delle informazioni presenti. Le operazioni sono state eseguite utilizzando il software ENVI. Ad ogni operazione è stato associato un *layer* di mascheramento relativo a:

- 1) Ombre,
- 2) Vegetazione .

La finalità del *layer* "ombre" è quella di individuare le aree in ombra a ridosso degli edifici.

In questo contesto individuare le zone in ombra equivale a determinare aree "*no-building*", ossia zone dove è esclusa la presenza di fabbricati. Piccole zone in ombra dovute alla presenza di torrioni scale, alberi, linee elettriche sono associate, invece, ad aree illuminate. L'ortofoto, per individuare le zone in ombra, è stata convertita dal sistema convenzionale RGB allo spazio colore HSV (*Hue, Saturation e Value*), in analogia a metodologie proposte anche da altri autori (Salvador, E. et al 2001, Tsai 2006). Per la creazione del *layer* "ombre" è stato eseguito il prodotto di due maschere, la prima sulla componente saturazione con valore minimo pari a 0.5, e la seconda sull'intensità con valore minimo pari a 0.5. Infine sono state eliminate tutte le zone in ombra aventi una superficie minore di 8 m² (circa 200 pixel).

La maschera ha numerosi ed estesi falsi positivi, ossia zone non in ombra associate come tali: corpi idrici profondi, alcune tipologie asfalto, automobili, alcune varietà di vegetazione, tutte entità che hanno *digital number* bassi nella banda del blu, non sono influenti nella metodologia proposta per l'individuazione ed estrazione degli edifici.

L'obiettivo del *layer* "vegetazione" è quello di distinguere ciò che è biomassa da ciò che non lo è, utilizzando i contenuti spettrali di un dato RGB. In questo studio, non disponendo dell'informazione della banda dell'infrarosso è stato deciso di combinare le bande del verde e del

rosso in modo analogo ai lavori di ricerca di Sibiriyakov (1996). La banda dell'infrarosso viene semplicemente sostituita con la banda del verde per formare l'NDGRI (indice normalizzato della differenza della banda del verde e del rosso).

Per la creazione della maschera NDGRI è stato impostato un valore di soglia minimo pari a 0.02, tale valore è stato ricavato empiricamente dall'analisi delle firme spettrali delle diverse entità.

Un limite dell'indice NDGRI è legato ad una erronea classificazione per alcune tipologie di tetti.

Per limitare questi effetti è stato introdotto il parametro di tessitura "*homogeneity*" della matrice co-occorrenza, definita anche GLCM (*Grey Level Co-Occurrence Matrix*). Per la creazione della maschera "*homogeneity*" è stato impostato un valore di soglia massimo pari a 0.5, ricavato empiricamente dall'analisi dei valori corrispondenti alle biomasse presenti.

Le informazioni spettrali dei due *layer* sono state combinate realizzando un'intersezione e selezionando solo le aree aventi una superficie maggiore di 4 m² (circa 100 pixel).

Infine è stato applicato l'operatore morfologico "dilatazione" con un *kernel* quadrato di lato 5 pixel, per ridurre la presenza di "buchi" nei poligoni e contenere il numero dei vertici di ogni area.

La quarta fase della procedura prevede la creazione del nDSM. In questo studio con il termine nDSM si indica un modello digitale di superficie "normalizzato" di punti "staccati" dal terreno. Tale modello si ottiene sottraendo i valori in quota del DSM a quelli del DEM. In questo modello i punti adagiati sul terreno hanno valori prossimi allo zero, mentre ad esempio alberi ed edifici, ecc., hanno valori maggiori di zero.

La procedura per la determinazione del nDSM può essere divisa in 3 parti. La prima parte si basa sulla determinazione di punti 3D, alla stessa stregua di quanto già fatto nella prima fase della procedura proposta, utilizzando, invece, una strategia di ricerca punti diversa da "*high urban*" e confrontando poi i *mass point* di entrambe le strategie.

Tale metodologia si ispira ad un lavoro proposto da Gooch et al. nel 2001.

La seconda nuvola di punti è stata generata utilizzando la strategia "*low urban*". L'idea di base è quella che se un punto per le entrambe le strategie fornisce lo stesso valore di quota, considerando ovviamente un piccolo scarto, quel punto con molta probabilità è un punto sul terreno.

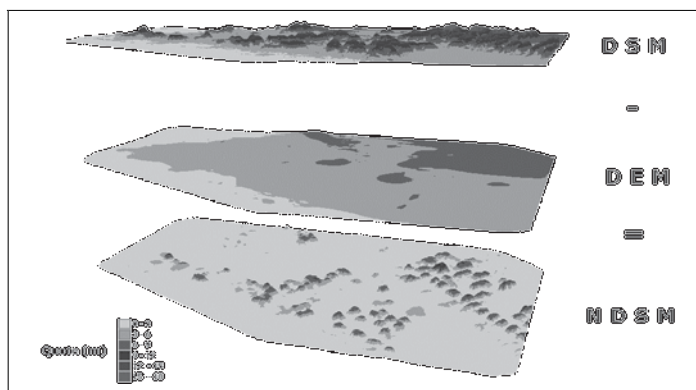


Figura 3 – I tre diversi dati altimetrici

E' stata eseguita, di seguito, una segmentazione della differenza dei due modelli altimetrici, impostando una soglia di taglio pari a 0.05 cm, una metrica city-block per i contorni, e un numero di pixel per ogni area non inferiore a 200. Nella seconda parte della procedura nDSM è stato eseguito il filtraggio per ottenere punti adagiati sul terreno, in particolare sono stati selezionati con una query spaziale soltanto quei punti contenuti nelle aree estratte nella fase precedente. Nella terza parte, infine, è stata eseguita la differenza fra i punti prodotti con la strategia "*high urban*" e i punti filtrati nella procedura proposta per ottenere il modello nDSM (Fig. 3).

La quinta e ultima fase della procedura è quella più sensibile ai parametri introdotti che, in funzione delle condizioni presenti nella scena acquisita, possono subire modifiche e integrazioni. In questa fase tutte le informazioni ricavate in precedenza vengono utilizzate. I parametri impostati dipendono dalla conoscenza a priori dell'area (dimensioni ed altezza media edifici, etc.) da parte dell'operatore. Nella quinta fase le informazioni presenti nei *layer* spettrali (Vegetazione e Ombre), sono state utilizzate come maschere per il *layer* altimetrico (nDSM).

Successivamente è stata effettuata una ulteriore segmentazione impostando una soglia minima pari a 2,5 metri ed una dimensione minima dei segmenti pari a 1000 pixel ($\sim 40 \text{ m}^2$).

I punti omologhi, determinati nella seconda fase e contenuti nei segmenti ultimi, sono stati utilizzati per creare triangoli con procedure di “*triangulation/merge/dissolve/envelope*” in ambiente ESRI ArcGis. Edifici erroneamente individuati sono stati eliminati mediante apposite “*query*” sulla dimensione ($\text{area} > 20 \text{ m}^2$) e sulla forma dei poligoni. Nella figura 4 è rappresentato il risultato finale dell'intera procedura.



Figura 4 – Edifici individuati automaticamente (in campitura) confrontati con restituzione classica (colore grigio)

Risultati e conclusioni

I risultati sono stati confrontati utilizzando gli indici DP (*detection percetage*) DP e il BF (*branch factor*) (Lin e Nevatia, 1998) sotto riportati:

$$DP = \frac{100 * TP}{TP + TN} \quad [1]$$

$$BF = \frac{100 * FP}{TP + FP} \quad [2]$$

Gli indici sono stati calcolati confrontando i risultati della procedura proposta con gli analoghi ottenuti attraverso restituzione fotogrammetrica classica, dove TP (*true positive*) rappresenta un edificio individuato correttamente sia dal restitutorista che dalla procedura automatica, FP (*false positive*) è un edificio riconosciuto solo dall'approccio automatico ma non dall'operatore, TN (*true negative*) indica infine un edificio distinto solo dall'esperto ma non dalla procedura automatica (Fig 4).

L'indice DP descrive il numero in percentuale di edifici correttamente riconosciuti dalla procedura automatica, mentre l'indice BF, al contrario, esprime il numero degli edifici individuati erroneamente.

Nella figura 5 sono evidenziati i risultati della metodologia proposta considerando 53 edifici presenti nell'area test:

TP	FP	TN	DP	BF
45	4	5	90.00%	8.16%
<i>Figura 5- Risultati del Test</i>				

Il metodo proposto ha dato “buoni” risultati se confrontati con altre metodologie proposte da diversi autori. La procedura automatica di estrazione per alcuni edifici risente ancora degli errori di *matching* generati nella seconda e quarta fase, legati alla presenza di occlusioni, vegetazione, contenuto spettrale simile tra coperture e suolo, ecc..

Le differenti tipologie di strutture presenti nell'area di prova consentono di affermare che, in aree con una maggiore uniformità edilizia, la metodologia proposta garantirà risultati migliori.

Riferimenti bibliografici

- Baltsavias E.P., Gruen A., Gool, (2001): “*Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*”, A.A. Balkema Publishers.
- Gooch, M. J., J. H. Chandler (2001): “*Failure prediction in automatically generated digital elevation models*”, Computers & Geosciences, 27(8): 913-920.
- Guelch E. (2000): “*Digital Systems for Automated Cartographic Feature Extraction*”, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B2: 241-256.
- Lin C., Nevatia, R. (1998): “*Building detection and description from a single intensity image*”, Computer Vision An image Understanding, CVIU 72(2), 101-121.
- 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, 1545-1548.
- 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): 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, 38-49.