

ESTRAZIONE AUTOMATICA DI OGGETTI DA IMMAGINI NEL SETTORE DEI BENI CULTURALI: PROBLEMATICHE ED APPLICAZIONI

Gabriele BITELLI, Valentina GIRELLI, Luca GUSELLA, Antonio ZANUTTA

DISTART, Viale Risorgimento 2, 40136 Bologna - ITALIA
(gabriele.bitelli, valentina.girelli, luca.gusella, antonio.zanutta@mail.ing.unibo.it)

Riassunto

Nell'ambito delle applicazioni relative al rilievo e alla classificazione dei beni culturali, la fotogrammetria digitale e la produzione di ortofoto sono una tecnologia di crescente interesse e di utilizzo diffuso. L'integrazione tra informazione *raster* e disegno vettoriale per la rappresentazione dell'architettura consente di evitare in parte la soggettività della restituzione al tratto, integrando le informazioni grafiche con utili particolari qualitativi. Negli ultimi anni, le tecnologie di processamento delle immagini digitali si sono dirette, partendo dall'uso del dato *raster* nel supporto alla digitalizzazione assistita, anche all'estrazione semi-automatica ed automatica di veri e propri oggetti di interesse dalle immagini. L'applicazione di tecniche di classificazione del tipo *object-oriented*, in rapido sviluppo per il dato telerilevato da satellite, pur avendo visto alcune realizzazioni nella classificazione di materiali costruttivi (marmi, pietre ornamentali, inerti), è ancora per certi versi in una fase iniziale nell'ambito dell'elaborazione di immagini di tipo architettonico o archeologico, e rappresenta in ogni caso un utile esercizio di ricerca. Il presente lavoro si propone di valutare le opportunità offerte dalle tecniche di estrazione automatica e assistita di oggetti su casi diversi, in particolare utilizzando orto-immagini relative ad opere di interesse storico ed architettonico.

Abstract

The object-oriented approach is now widely applied in the analysis of remote sensing imagery and aerial Photogrammetry. A few experiences are instead developed for the analysis of close-range photogrammetric products, like orthophotos and photoplanes, but in the survey of Cultural Heritage some applications could be of interest. The work shows some preliminary results in this direction, based on simple study cases.

1.Introduzione

Il settore del rilevamento dei Beni Culturali ha visto in questi anni un forte ampliamento delle possibilità di acquisizione e rappresentazione dei dati, con lo sviluppo di nuove tecniche e tecnologie che offrono soluzioni competitive rispetto a quelle tradizionali in termini di produttività, di qualità, di integrazione e di flessibilità di uso dei prodotti derivati. Se rimane indubbiamente cruciale l'interazione e l'intervento diretto dell'esperto, sia in fase di misura che di interpretazione del dato, la disponibilità di prodotti in forma digitale può consentire nuovi approcci e favorire momenti di automazione, parziale o totale, nella gestione delle informazioni primarie o derivate.

Tra gli aspetti che sono stati oggetto di sperimentazione e studio vi è quello della estrazione di caratteristiche, sia di carattere semantico che più semplicemente di tipo geometrico, da immagini *raster*. La struttura matriciale di un'immagine, con la sua organizzazione sistematica in *pixel* dotati di uno o più attributi, relativi a colore o altre variabili, si presta all'applicazione di algoritmi automatici, che in genere si basano sui valori numerici dei singoli *pixel*; un approccio diverso è invece quello che tende a raggruppare i *pixel* in "oggetti" autonomi (segmentazione), per poi procedere ad una classificazione operando direttamente su questi ultimi. Questo tipo di soluzione,

che sarà più ampiamente descritta più oltre riferendosi ad una specifica implementazione software in commercio, viene definita “orientata agli oggetti”. La sua natura potrebbe apparire più consona a dare risposta alle richieste di interpretazione di immagine che vengono dal campo dei beni culturali, per esempio architettura ed archeologia, quando gli operatori ed i ricercatori del settore necessitano di estrarre contenuto informativo – di natura qualitativa o anche quantitativa – da immagini fotografiche, che nel caso di ortofoto sono dotate anche di una valenza metrica. Non ci si riferisce con ciò necessariamente all’esecuzione di una vera e propria restituzione vettoriale al tratto, che ancora resta obiettivo di difficile raggiungimento in modalità automatica, come del resto accade per i fotogrammi aerei, ma ad altre funzionalità che potrebbero essere di interesse nello studio del costruito, per esempio nel progetto e gestione di operazioni di restauro (valutazione della estensione di aree con certe caratteristiche in una facciata, studio sulla distribuzione delle tessere e delle lacune in una superficie musiva, ecc.), nella predisposizione di un’analisi strutturale in cui occorra definire sommariamente la forma e la disposizione dei principali elementi costruttivi, ecc.

Quello che si presenta è l’inizio di un tentativo di applicare alle immagini derivanti dalla fotogrammetria *close-range* delle metodologie che sono in corso di studio per l’analisi di immagini telerilevate da satellite o da aereo (Bitelli et al., 2005), senza alcuna pretesa di completezza ma con il solo intento di valutare, o meglio iniziare a valutare, la fattibilità di nuovi percorsi. Di seguito verrà presentato per sommi capi l’approccio *object-oriented*; seguiranno due casi di studio con immagini che sono state sottoposte a segmentazione – che costituisce evidentemente la fase cruciale - e successiva classificazione, utilizzando il software eCognition sviluppato da Definiens Imaging, che appare tra i più interessanti prodotti di questo tipo (Meinel & Neubert, 2004).

2. La segmentazione di immagini in un approccio object-oriented

La segmentazione è la suddivisione di una immagine in sezioni separate. L’approccio di segmentazione multi-risoluzione di eCognition (Baatz & Schäpe, 2000) è di tipo *bottom-up*, ovvero un sistema di unione delle regioni contigue a partire da ogni singolo pixel. Due oggetti all’interno di un’immagine sono uniti insieme quando un criterio di eterogeneità è raggiunto. Questo criterio di eterogeneità, o scala, è essenzialmente una misura della magnitudine di un oggetto. La segmentazione in eCognition è di tipo multi-risoluzione, ovvero possono convivere all’interno della struttura dati segmentazioni di dimensioni diverse, in cui sono definiti dei limiti tra oggetto e oggetto in senso orizzontale e dei vincoli in senso verticale.

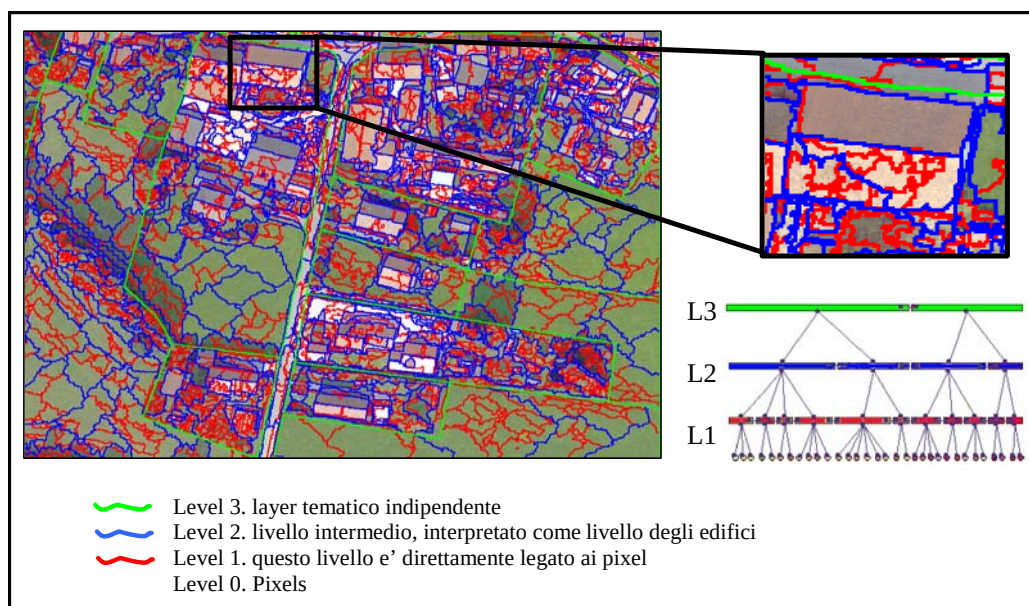


Figura 1 - Segmentazione multi-risoluzione, con indicate le relazioni tra livelli. Il livello più basso (L1) è in relazione direttamente con i *pixel*. Da qui verso l’alto per il livello 3. Il livello 3 è invece un *layer* tematico indipendente (per es. per un dato territoriale uno *shapefile* di confini catastali).

Nella figura 1, relativa ad un'immagine aerea, sono riportate le gerarchie di livelli in eCognition: i *pixel* sono nel livello più basso (L0). Gli oggetti nel livello 1 (L1) contengono tutti gli oggetti del livello 0 e hanno il vincolo di essere contenuti (e quindi avere i lati in comune) negli oggetti del livello superiore. Identicamente, il livello 2 contiene tutti gli oggetti del livello 1 ed è vincolato dal livello 3.

Come tecnica *bottom-up*, la segmentazione comincia con un singolo *pixel*, nel livello L0. Due oggetti sono fusi insieme quando un dato valore di differenza di eterogeneità (*Scale, f*) è raggiunto tra gli oggetti separati ed uniti.

$$Scale(f) = w_{color} \cdot h_{spectral} + (1 - w_{color}) \cdot h_{shape} \quad [1]$$

Questo valore di scala è una media pesata (w_{color} rappresenta il peso) tra un fattore di eterogeneità spettrale ($h_{spectral}$) ed un valore di eterogeneità di forma (h_{shape}). Un valore di peso w_{color} pari ad 1 significa utilizzare solamente l'eterogeneità spettrale, mentre il valore 0.1 è il minimo valore utilizzabile per l'eterogeneità spettrale e massimo per quella di forma.

Il fattore spettrale (o di colore) è definito come la deviazione standard (σ_c) dei *pixel* costituenti un oggetto, considerando le differenti bande utilizzate tramite un fattore di peso (w_c): quest'ultimo permette di utilizzare dati di differenti risoluzioni spaziali e diversi *bit depth*, consentendo inoltre di usare tipi diversi di immagini *raster* (ad es. un DEM insieme a immagini telerilevate multispettrali).

$$h_{spectral} = \sum w_c \cdot \sigma_c \quad [2]$$

L'omogeneità di forma (3a) è a sua volta una media pesata, attraverso il peso ($w_{compactness}$), tra un fattore di compattezza ($h_{compact}$) e un fattore di *smoothness* (h_{smooth}). Il fattore di compattezza (3b) è il rapporto tra la lunghezza del perimetro (l) e la radice quadrata del numero di *pixel* compreso nell'oggetto (n). Il fattore di *smoothness* (3c) è invece il rapporto tra il perimetro dell'oggetto e il perimetro del *bounding box* (b) che lo contiene.

$$h_{shape} = w_{compactness} \cdot h_{compact} + (1 - w_{compactness}) \cdot h_{smooth} \quad [3a]$$

$$h_{compact} = l / \sqrt{n} \quad [3b]$$

$$h_{smooth} = l / b \quad [3c]$$

Riassumendo, i parametri che possono influenzare la segmentazione sono diversi:

- *bit depth*: il fattore di eterogeneità spaziale è calcolato sulla deviazione standard di un'immagine, per questo motivo l'uso di immagini a 8 oppure a 16 bit è rilevante;
- pre-processamento dell'immagine: equalizzazione del contrasto, grado di compressione, formato, per le stesse ragioni esaminate sopra;
- bande usate: sono possibili diverse combinazioni di bande e diversi pesi; ad esempio, per le immagini satellitari ad alta risoluzione è possibile utilizzare la banda pancromatica separatamente con le 4 bande multispettrali, oppure utilizzare direttamente i prodotti di *data fusion*;
- parametri di segmentazione.

3. La segmentazione di ortofoto e fotopiani: alcuni casi di studio

Per la sperimentazione in corso, che si intende sviluppare ed ampliare nel futuro, sono state scelte immagini legate a recenti attività di rilevamento condotte dal DISTART nel settore dei Beni Culturali e relative a varie scale di lavoro.

Un primo esempio si riferisce al particolare di una ortofoto relativa ad una superficie musiva del Ninfeo della "Domus dei Gemelli" in Pompei, realizzata nell'ambito del rilevamento sistematico dei principali ninfei pompeiani (Bitelli et al., 2002), mentre un secondo è costituito dal fotopiano di una parete dei bastioni di Floriana, in Malta (Baratin et al., 2001).

La tipologia degli oggetti di interesse è evidentemente diversa. Nel primo caso potrebbe essere di interesse il riconoscimento automatico delle singole tessere del mosaico, il loro conteggio ed un inventario basato su caratteristiche cromatiche o materiche, la valutazione dell'entità delle lacune, ecc., per un utilizzo che potrebbe essere legato ad una fase di restauro. Nel secondo caso si vuole estrarre il perimetro dei blocchi che costituiscono la struttura della muratura e che descrivono l'arco, ad esempio per un primo supporto ad una analisi strutturale.

Una prima sperimentazione ha evidenziato che vi sono notevoli differenze nei risultati della segmentazione nei due casi.

Considerando l'ortofoto a colori del mosaico pompeiano, la sperimentazione con l'approccio ad oggetti è avvenuta considerando la specifica tipologia degli oggetti di interesse (le tessere del mosaico, di diversi materiali e colore) e la loro relazione con i parametri di segmentazione.

Nell'esempio di figura 2 (relativo ad una piccola porzione del mosaico) sono riportati gli effetti del parametri w_{color} , indicante la quantità relativa all'equazione [2], rispetto a quella relativa alle equazioni [3], a parità del parametro scale (40, equazione [1]) e del peso del fattore di forma ($w_{compactness}$), posto pari ad 1. L'utilizzo di un fattore di compattezza elevato è giustificato dalla caratteristica degli oggetti di estrarre.

Diminuendo la quantità di colore da inserire nella segmentazione, vengono delineati i contorni dei singoli grani del mosaico, indipendentemente dalla loro eterogeneità interna. E' quindi possibile notare che, se pure le tessere contengono difetti o scheggiature, tali imprecisioni sono correttamente comprese nell'oggetto delineato.

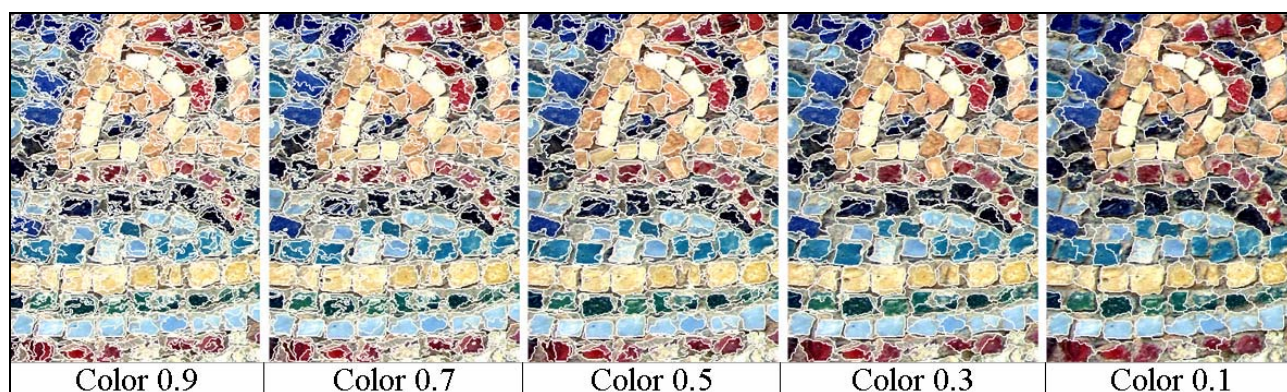


Figura 2 - Influenza del parametro Color a parità di scala (40) e fattore di forma (0.5).

Fissando il parametro Color 0.3, si è quindi valutato il parametro scala, considerando che, per essere correttamente riconosciuti, i grani devono essere distinguibili gli uni dagli altri. La figura 3 che segue è esplicativa: l'analisi è cominciata dalla scala 20 fino alla scala 60, quando gli oggetti sono divenuti troppo grandi per comprendere univocamente una sola tessera di mosaico.

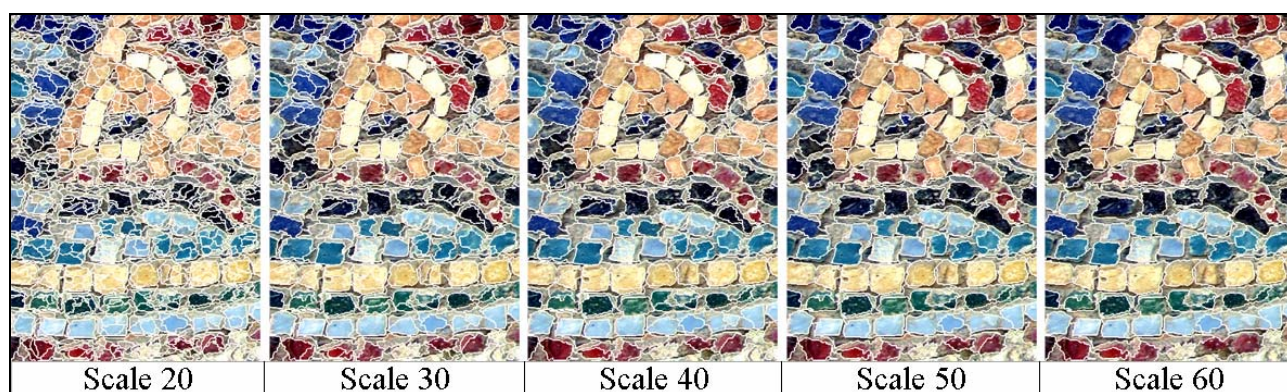


Figura 3 - Influenza del parametro scala.

Si sono evidenziate due tipologie di problematiche:

- 1) alcune tessere di mosaico sono costituite da più di un oggetto (per valori elevati di scala)
- 2) alcuni oggetti comprendono al loro interno più di una tessera mosaico (per valori bassi di scala).

I parametri ottimali della segmentazione saranno raggiunti quando saranno minimizzati entrambi gli aspetti negativi. Come si può notare dalla Figura 4, il maggior numero di tessere corrette si trova per il valore di scala 50 (problematica 1), mentre il numero di oggetti che intersecano non correttamente le tessere digitalizzate tende a crescere all'aumentare della scala (problematica 2).

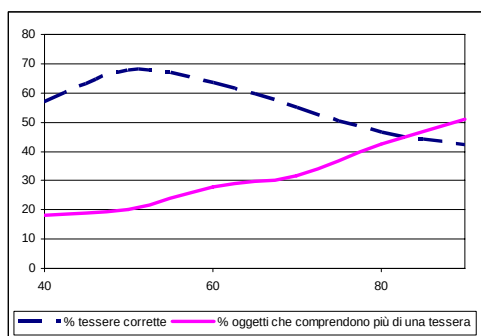
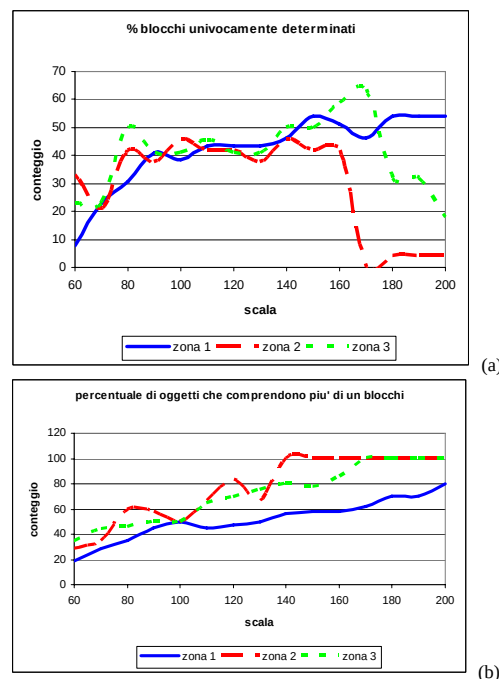


Figura 4 - Ottimizzazione dei parametri di segmentazione delle tessere.

Il secondo caso ha fornito invece risultati non soddisfacenti, o perlomeno ha evidenziato la necessità di parametrizzare la segmentazione in modo differente per le diverse tipologie di oggetti di interesse presenti nell'immagine, che d'altra parte non costituisce un caso semplice: come si può osservare dalla figura 5, su questa porzione di mura vi sono infatti tre diverse tipologie di blocchi, che si differenziano per caratteristiche cromatiche, di materiale e anche disposizione.



(a)

(b)

Figura 5 – Fotopiano di muratura nel bastione di Floriana, Malta: esempio di segmentazione con un certo set di parametri (colore 0.3, compattezza 1.0) e valutazione dei risultati su tre zone diverse, riconoscibili dal colore di una schematica vettorializzazione manuale. I grafici (a) e (b) confermano come, al crescere del fattore di scala, la sola zona 1 (blu) fornisca risultati abbastanza buoni.

La figura 5 mostra una delle prove effettuate, per le quali si è adottato un parametro colore pari a 0.3 ed un parametro di compattezza pari ad 1; nella figura è riportata con colori una vettorializzazione manuale grezza di riferimento e con tratto nero il risultato della segmentazione. Come si può osservare, e la conclusione si potrebbe d'altra parte riferire ad una qualsiasi delle molte prove effettuate, i parametri sono utili ad individuare abbastanza bene solamente una zona, per la quale i conci vengono scontornati con discreta precisione, mentre nelle altre il risultato è del tutto insoddisfacente. Ciò può essere confermato anche graficando alcune variabili significative (figura 5, (a) e (b)), quali la percentuale di conci univocamente determinati e quelli di oggetti che, invece, comprendono più di un blocco; i grafici sintetizzano il risultato di molteplici esperimenti condotti variando il fattore di scala, riportato in ascissa. Si può notare che al crescere di tale fattore nelle zone 2 e 3 sale molto rapidamente la percentuale di oggetti che inglobano più di un concio, e discende molto rapidamente quella dei blocchi determinati in modo univoco. Su immagini di questo tipo non è dunque semplice effettuare la segmentazione in modo soddisfacente, o quantomeno essa dovrebbe essere eseguita più volte con i parametri migliori per ciascuna categoria di oggetti di interesse.

4. Classificazione

Una volta determinati i parametri di segmentazione, si può passare alla fase di classificazione, che qui si mostra per il caso di studio del mosaico. Il sistema di classificazione scelto è supervisionato, utilizzando una classe per ciascun colore di tessera presente; per migliorare la classificazione, sono inoltre utilizzate le caratteristiche geometriche degli oggetti. Da un'analisi visuale delle tessere, appare chiaro che si è interessati alla classificazione di oggetti più grandi di una certa area, non allungati, e con poche ramificazioni: queste caratteristiche sono riportate in eCognition come una serie di *threshold*, permettendo una estrema pulizia della classificazione finale. La figura 6 mostra le fasi ed i risultati conseguiti.

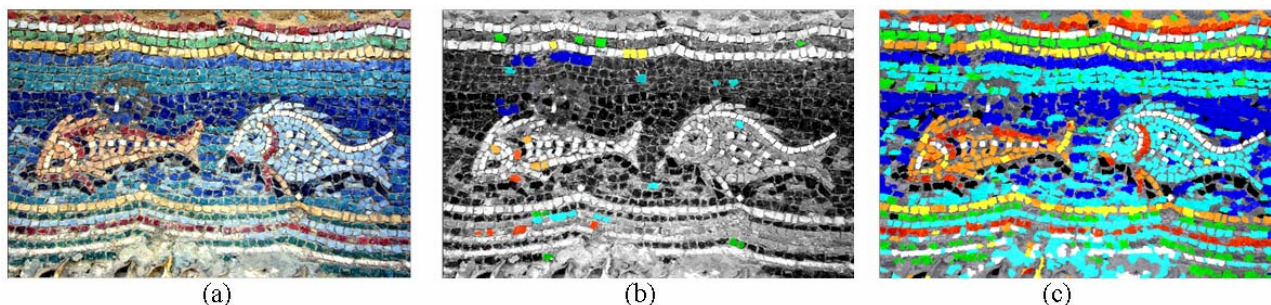


Figura 6 – (a) immagine originale (b) *sample* selezionati (c) risultato della classificazione

E' possibile pervenire così ad un conteggio delle tessere secondo il loro colore, contando gli oggetti classificati; il risultato è riportato nella tabella 1.

Conclusione

La sperimentazione, per quanto ancora in una fase iniziale, mostra le potenzialità dell'utilizzo delle tecniche di classificazione object-oriented nello studio dei Beni Culturali. Le principali problematiche emerse derivano dalla definizione dell'oggetto di interesse, dalla parametrizzazione della procedura di segmentazione e dalla caratterizzazione dei *sample* per la classificazione, nonché infine dalla valutazione dell'accuratezza; su questi temi proseguirà la ricerca.

colore	n.ro oggetti
rosso	162
bianco	146
azzurro	570
blu	380
verde	166
giallo	95
ocra	154
nero	139

Tabella 1 – Conteggio automatico del numero di tessere a diverso colore

Bibliografia

- Baatz M, Schäpe A. (2000), "Multiresolution segmentation – an optimisation approach for high quality multi-scale image segmentation", *Angewandte Geographische Informationsverarbeitung XI, Vorträge zum AGIT-Symposium Salzburg*, Strobl J, Blaschke Th, Griesebner G (eds). Herbert Wichmann, Heidelberg, 12-23.
- Baratin L., Bitelli G., Bonnici H., Unguendoli M., Zanutta A. (2001), "Traditional and modern methods of surveying architectural heritage: a few examples in the fortified island of Malta". *Proc. CIPA2001*, Berlino, 529-535.
- Benz, U., Hofmann, P., Willhauck, G., Lingenfelder, I. and M. Heynen (2004), "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information", *ISPRS Journal of Photogrammetry & Remote Sensing* 58 (2004), 239-258.
- Bitelli G., Capra A., Zanutta A. (2002), "Digital Photogrammetry and Laser Scanning in Surveying the "Nymphaea" in Pompeii". *Proc. CIPA WG6 Int. Workshop on Scanning for Cultural Heritage Recording*, Corfu, 115-120.
- Bitelli G., Gusella L., Mognol A., (2005), "Determinazione dei parametri di segmentazione in immagini aeree e telerilevate: un approccio semantico all'ottimizzazione", *Convegno SIFET 2005*, Palermo
- Meinel G., Neubert M, (2004), "A comparison of segmentation programs for high resolution remote sensing data", *ISPRS Archives*, Istanbul.