

Mosaicatura e georeferenziazione di immagini iperspettrali

Antonio Cenedese (*), Carlo Dacquino (**), Monica Moroni (*)

(*) DICEA, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome (Italy),
+390644585095, +390644585094, monica.moroni@uniroma1.it; antonio.cenedese@uniroma1.it

(**) ISPRA, Istituto Superiore per la Protezione e la Ricerca Ambientale,
Via del Curtadone 3, 00185 Rome (Italy), carlo.dacquino@ispambiente.it

Riassunto

Presso il Dipartimento di Ingegneria Civile Edile e Ambientale della Sapienza Università di Roma sono stati messi a punto due sistemi di acquisizioni di immagini iperspettrali (caratterizzati da basso costo, peso e consumo energetico e quindi facilmente trasportabili) basati sull'utilizzazione di filtri interferenziali tunabili e di spettrometri. È stato inoltre sviluppato un robusto ed efficiente algoritmo per mosaicare in modo automatico, senza la necessità di individuare punti di riferimento, molte immagini tra loro parzialmente sovrapposte in modo da ottenere un'unica scena del territorio monitorato. L'algoritmo è basato sull'analisi della correlazione, valutata tramite FFT 2D, esistente tra due immagini successive, e consente di valutare lo spostamento relativo tra le immagini stesse tenendo conto della rotazione e del cambiamento di scala. Le immagini sono inoltre georeferenziate nel sistema di coordinate geografiche WGS-84. Le prestazioni della procedura sono state valutate utilizzando le immagini, riprese da un aeromobile superleggero, in una campagna proximal sensing svolta nella piana di San Teodoro in provincia di Olbia avente come scopo la caratterizzazione del terreno attraverso il suo cubo iperspettrale. Le mappe che mettono in evidenza differenti oggetti sono in ottimo accordo con i rilievi *ground-truth* effettuati.

Abstract

At the Department of Civil and Environmental Engineering of Sapienza University of Rome two systems for the acquisition of hyperspectral images (characterized by low cost, weight and power consumption) based on the use of interference filters and spectrometers were developed. A robust and efficient algorithm for automatic image mosaicing, without the need to identify reference points, was developed. The algorithm, together with the 2D fast Fourier transform, provides an estimate of the displacement between pairs of images by accounting for rotations and changes of scale. The resulting mosaic was successively georeferenced within the WGS-84 geographic coordinate system. The performances of the algorithm were evaluated using sample images and image sequences acquired during a proximal sensing field campaign conducted in San Teodoro (Olbia-Tempio—Sardinia). The hyperspectral cube closely corresponds to the mosaic. Mapping allows for the identification of objects within the image and agrees well with ground-truth measurements.

Introduzione

Le grandi potenzialità del monitoraggio attraverso l'analisi di immagini iperspettrali sono ampiamente documentate nella più recente letteratura scientifica trovando applicazione in differenti settori, tra cui:

- il monitoraggio del suolo per evidenziare possibili situazioni di degrado che possono dar luogo a processi di desertificazione (Stocking and Clark 1999);
- la classificazione della vegetazione che consente di identificare differenti coltivazioni e l'insorgere di situazioni di stress (Sims and Gamon 2002);

- il monitoraggio dei beni culturali per identificare i materiali utilizzati e lo stato di degrado delle singole opere al fine di progettare gli interventi necessari per la loro messa in sicurezza (Tapete *et al.* 2012);
- la determinazione del livello di inquinamento e di trofia del mare e dei corpi idrici, in particolare il monitoraggio della morfologia di ambienti lagunari e della distribuzione della vegetazione alofila negli ambienti intertidali (Lillesand *et al.* 2004);
- le applicazioni in campo idrologico al fine di valutare l'evapotraspirazione, di difficile misurazione diretta (Minacapilli *et al.* 2012);
- la separazione di sostanze, in particolare dei rifiuti, nelle differenti tipologie di materiali (Tatzer *et al.* 2005).

Nonostante le sue indiscusse potenzialità, il monitoraggio basato sull'analisi delle immagini iperspettrali non è oggi molto diffuso. Utilizzando i più recenti sviluppi hardware e software nell'acquisizione ed elaborazione delle immagini sono state aumentate considerevolmente le prestazioni della tecnica in termini di risoluzione spaziale e temporale (Moroni and Cenedese, 2012). Permangono tre problematiche che ne limitano l'utilizzo, con riferimento ai sistemi sviluppati fino a qualche anno fa (ad esempio il sistema MIVIS, Multispectral Infrared and Visible Imaging Sensor, acquisito dal CNR nel 1994 nell'ambito del progetto LARA (Bianchi *et al.* 1997)):

- il costo di acquisizione e di manutenzione dei sistemi è molto elevato;
- i sistemi sono difficilmente trasportabili e richiedono una elevata disponibilità di energia per il loro funzionamento;
- anche utilizzando il software commerciale, l'elaborazione delle immagini iperspettrali per ottenere mappe tematiche risulta molto onerosa; in particolare, la gestione delle immagini, che va dall'acquisizione alla loro interpretazione per arrivare ad un risultato di facile utilizzazione, richiede l'intervento di persone molto specializzate.

Uno dei problemi ancor oggi oggetto di approfondimento è quello della mosaicatura automatica di sequenze di immagini che permette la ricostruzione completa della scena che si intende analizzare. In alcuni casi si tratta di sequenze di migliaia di immagini, pertanto appare impraticabile il riconoscimento di punti corrispondenti nelle diverse immagini. In precedenti lavori (Moroni *et al.* 2012), facendo riferimento ad una trasformazione omografica semplificata, è stato fatto dell'uso iterativo della funzione di correlazione bidimensionale tra due immagini successive per valutare il loro cambiamento di scala, rotazione e spostamento. Partendo da un'immagine di riferimento, la scena veniva ricostruita correlando via via le immagini successive. Tuttavia, anche se la prima immagine è georeferenziata, gli errori progressivamente ed inevitabilmente introdotti comportano il fatto che le immagini successive non siano più correttamente georeferenziate. Per ovviare a questo inconveniente nel presente lavoro si propone di correlare le immagini non tra di loro ma con quelle di una mappa di riferimento correttamente georeferenziata, ad esempio Google Earth.

Presso il Dipartimento di Ingegneria Civile Edile e Ambientale (DICEA) negli ultimi anni sono stati messi a punto due sistemi di acquisizione delle immagini che hanno permesso di superare gran parte dei limiti ricordati. I due sistemi, uno basato sull'uso di spettrometri l'altro basato sull'uso di filtri interferenziali, utilizzano le più recenti tecnologie di acquisizione di immagini ad alta risoluzione ed alta frequenza e coprono le lunghezze d'onda comprese tra 400 nm e 1800 nm, proprie del visibile (VIS) e del vicino infrarosso (NIR) (Cenedese *et al.* 2007). I sistemi sono completati con una telecamera termica in grado di rilevare la temperatura nei vari punti della scena ripresa. Gli apparati sviluppati, di basso costo, peso e consumo energetico, si prestano ad essere utilizzati per monitoraggio in *remote sensing* (a grande distanza, ad esempio da satellite, da cui non è possibile, in genere, raggiungere la scena ripresa) e *proximal sensing* (a piccola distanza, installati su aeromobili di piccole dimensioni, quando è possibile raggiungere la scena ripresa e verificare a cosa corrispondono gli spettri acquisiti attraverso operazioni di *ground truth*).

In seguito sarà descritto solo il sistema che utilizza gli spettrometri, tuttavia le procedure di mosaicatura sono utilizzabili anche per immagini acquisite con il sistema che utilizza i filtri.

Sistema di acquisizione delle immagini basato sugli spettrometri

In questo sistema sono utilizzati due spettrometri, il primo centrato nel visibile (VIS da 400 nm a 1000 nm), il secondo centrato nel vicino infrarosso (NIR da 900 nm a 1800 nm). Con questo sistema si acquisiscono simultaneamente gli spettri relativi a punti equidistanti che si trovano su un segmento. È, quindi, necessario riprendere più immagini per ricostituire una scena bidimensionale attraverso l'accostamento di più segmenti.

Nella Figura 1 è riportato lo schema del sistema realizzato comprendente:

- due frame grabber ciascuno con due ingressi Camera Link;
- spettrometro VIS (S1) montato di fronte una telecamera CMOS;
- spettrometro NIR (S2) montato di fronte un telecamera InGaAs;
- telecamera CMOS;
- telecamera termica;
- sistema di alimentazione di tutti gli apparati;
- calcolatore di processo per controllare tutto il sistema ed acquisire tramite porta USB le immagini della telecamera termica.

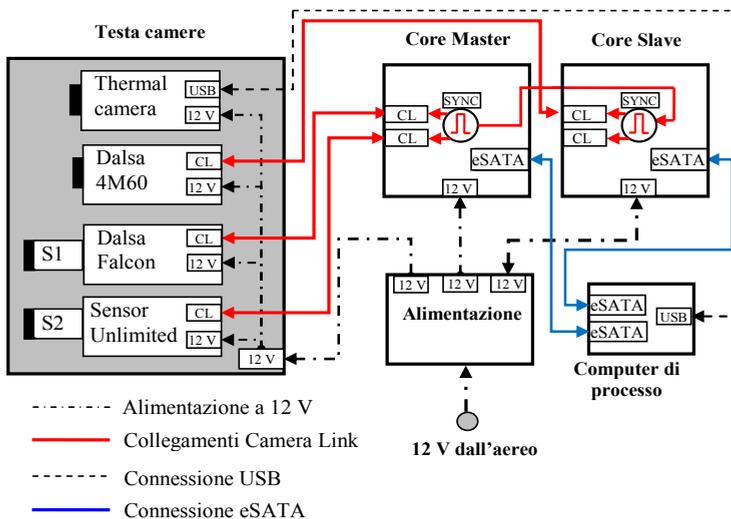


Figura 1. Sistema con spettrometri.

Davanti agli spettrometri sono montati due obiettivi da 25 mm, uno dei quali realizzato in modo da non assorbire la radiazione NIR.

Sono utilizzate due telecamere CMOS in bianco e nero: una Dalsa 4M60 dotata di un sensore di 4 Mpixels (2352×1728) quadrati (7.4 μm×7.4 μm), in grado di acquisire fino a 30 immagini al secondo a risoluzione piena ed un'uscita a 8 o 10 bit e una Dalsa Falcon 1.4M100 dotata di un sensore di 1 Mpixels (1280×1024) quadrati (7.4 μm×7.4 μm), in grado di acquisire fino a 100 immagini al secondo a risoluzione piena ed un'uscita a 8 o 10 bit. Delle due telecamere, la Dalsa Falcon è accoppiata allo spettrometro VIS, l'altra telecamera, la Dalsa 4M60, è utilizzata per acquisire un'immagine 2D e rendere possibile la corretta georeferenziazione nella scena della linea ripresa, correlando tra loro immagini successive.

Per acquisire nel NIR è stata utilizzata una telecamera InGaAs (nella quale il silicio costituente il sensore è drogato con arseniuro di indio e gallio) Sensor Unlimited SU320M-1.7RT. Il sensore è costituito da 320×240 pixels quadrati (40 μm×40 μm), in grado di acquisire fino a 50 immagini al secondo ed un'uscita a 8 o 10 bit. Tale telecamera si presta ad essere accoppiata al solo spettrometro NIR.

La telecamera termica Jade UC è equipaggiata con sensore microbolometrico non raffreddato di 384×288 pixels (con area sensibile pari a 320×240 pixels) quadrati (35 μm×35 μm) con risposta spettrale compresa nell'intervallo 8-12 μm; il segnale video è memorizzato in digitale a 14 bit.

Le immagini acquisite da tutte le telecamere utilizzate, salvo la termica, sono trasferite a 8 bit. Pertanto il livello di grigio, o *digital number* DN, di ciascun pixel è compreso tra 0 e 255.

Il sistema di acquisizione e memorizzazione delle immagini è costituito da due DVR Express ® Core della IO Industries in configurazione master-slave. Ciascun Core è dotato di due ingressi Camera Link in grado di trasferire fino a 780 Mbytes/s, la memoria è costituita da quattro dischi allo stato solido con una capacità di 1 Tbytes. Il segnale di trigger è generato dal Core master e sincronizza l'acquisizione di tutte le telecamere. Tale sistema permette di acquisire le immagini senza alcuna compressione.

Un calcolatore, tramite due porte eSATA, gestisce tutto il processo di sincronizzazione, ripresa e memorizzazione delle immagini. Il sistema è stato progettato in modo da ridurre i pesi (meno di 8 Kg) ed i consumi (meno di 300 W all'accensione).

Algoritmo di mosaicatura

Tutte le tecniche definite genericamente con il termine *frame fusion* hanno come obiettivo quello di riprodurre un'unica immagine a partire da una serie di più immagini ciascuna delle quali riporta tutta o solo una parte della scena. Si indica con *super resolution* il procedimento con cui si sovrappongono immagini della stessa scena al fine di migliorarne la qualità ed aumentare la risoluzione. Si indica con mosaicatura il procedimento con cui si mettono insieme più immagini parzialmente sovrapposte per ottenere una sola immagine della scena (Capel 2004).

La mosaicatura delle immagini si sviluppa schematicamente nelle operazioni di seguito descritte.

Allineamento (registration). L'obiettivo è di trasformare ogni immagine per inserirla in un unico riferimento di coordinate globali contenente l'intera scena. La scelta del riferimento globale non è rilevante, ma per comodità viene allineato con gli assi di una delle immagini della scena (generalmente la prima), l'immagine viene così chiamata "immagine di riferimento", in modo tale che la trasformazione tra il riferimento globale e l'immagine di riferimento sia l'identità.

Riproiezione (reprojection). Dopo l'allineamento tutti i punti di ogni immagine sono trasformati in punti dell'immagine globale finale attraverso operazioni di interpolazione.

Combinazione (blending). Una volta che tutte le immagini sono state allineate e riproiettate, ad ogni punto del riferimento globale corrispondono, in genere, differenti valori che vanno rielaborati al fine di giungere ad un solo valore che dia omogeneità fotometrica al mosaico finale. In questo modo si evita l'insorgere di vistose discontinuità nelle zone di sovrapposizione. Vengono utilizzate diverse funzioni per rendere omogeneo il mosaico, la più semplice delle quali è quella della media tra i valori delle immagini adiacenti.

Per effettuare la mosaicatura di più immagini è necessario utilizzare una trasformazione geometrica che definisce la relazione tra i punti di ciascuna coppia di immagini. Le trasformazioni possono essere globali o locali (Gonzales and Woods 1993). Le trasformazioni locali sono applicate solo ad una parte dell'immagine; per tali trasformazioni è molto difficile utilizzare una relazione generale concisa da poter applicare per tutte le riprese effettuate (Goshtasby 1986). Le trasformazioni globali possono essere gestite attraverso l'omografia quando le immagini sono relative a:

- superfici piane riprese da una telecamera in moto arbitrario;
- superfici non piane acquisite a grande distanza da una telecamera;
- superfici non piane acquisite da una telecamera in rotazione intorno al proprio asse ottico e/o in avvicinamento/allontanamento dal soggetto.

Sono necessari otto gradi di libertà perché l'omografia possa tener conto di traslazioni rigide, rotazioni rigide, variazioni di scala e scorrimenti tra due immagini (trasformazione proiettiva). Sono necessari sei gradi di libertà se il punto di ripresa è all'infinito (trasformazione affine). Sono necessari quattro gradi di libertà per una ripresa all'infinito se non si tiene conto dello scorrimento tra le immagini (trasformazione simile). Poiché operare attraverso una omografia completa è

piuttosto complesso, si suggerisce una metodologia meno rigorosa che si adatta tuttavia ad elaborare un numero elevato di immagini rilevate da un aeromobile in cui si suppone che non avvengano scorrimenti e che i cambiamenti di scala siano gli stessi nelle due direzioni.

La correlazione bidimensionale tra due immagini permette di valutare, attraverso la determinazione del suo picco, la traslazione rigida di un'immagine rispetto all'altra. L'algoritmo è basato sull'uso della trasformata di Fourier bidimensionale (2-DFFT).

Il posizionamento dell'*i*-esima immagine sulla scena individuata sulla mappa di riferimento georeferenziata è effettuata in due passi successivi:

- determinazione della variazione di scala, rotazione e spostamento, attraverso l'uso iterativo di correlazioni, con la *i-1* immagine estratta dalla mappa di riferimento; per tenere conto di rotazioni e cambiamenti di scala tra immagini consecutive, una immagine della coppia viene ruotata nell'intervallo -5° - 5° con passo 0.05° , e la scala viene modificata rispetto alla dimensione originale nell'intervallo -6% - 6% con passo 0.5% ; tra tutte le correlazioni effettuate si determina la rotazione ed il cambiamento di scala associati al massimo valore nel picco di correlazione
- estrazione dalla mappa di riferimento dell'*i*-esima immagine, utilizzando sempre correlazioni bidimensionali, da confrontare successivamente con la *i+1* immagine acquisita.

Correlazione tra immagine ripresa e Google Earth

Il punto più critico della procedura suggerita è quello che riguarda la correlazione tra immagini riprese, in bianco e nero, e quelle estratte da Google Earth a colori. Sono riportati di seguito alcuni risultati preliminari ottenuti utilizzando immagini riprese in una campagna di misura effettuata in località San Teodoro (Tempio-Olbia) con il sistema di acquisizione installato su di un aeromobile leggero che volando a circa 500 m di quota consente di ottenere una risoluzione a terra di circa 10 cm, molto più elevata di quella di Google Earth. Nelle Figura 2 e Figura 3 sono riportate un'immagine ripresa ed un particolare di Google Earth della stessa zona.



Figura 2. Immagine ripresa durante la campagna di misura effettuata in località San Teodoro (Tempio-Olbia).

Dopo un'opportuna rotazione, cambiamento di scala e trasformazione in bianco e nero si ottiene l'immagine di Google ricampionata (Figura 4) corrispondente a quella ripresa. Nella Figura 5 sono riportati i contributi puntuali alla covarianza (nero contributi negativi, bianco contributi positivi). Si osserva come, in genere, siano ben individuati i componenti antropici (edifici, piscine, ...) mentre le aree in grigio, legate alla vegetazione e alla illuminazione, risultino differenti nelle due immagini confrontate. Infine in Figura 6 è riportato il coefficiente di correlazione tra le immagini di Figura 2 e Figura 4. Pur essendo il massimo di tale coefficiente non molto elevato (circa 0.5), la presenza di un picco molto ben individuato nell'origine permette di collocare correttamente l'immagine ripresa nella mappa di riferimento georeferenziata.



Figura 3. Particolare della mappa Google Earth della stessa zona della Figura 2.

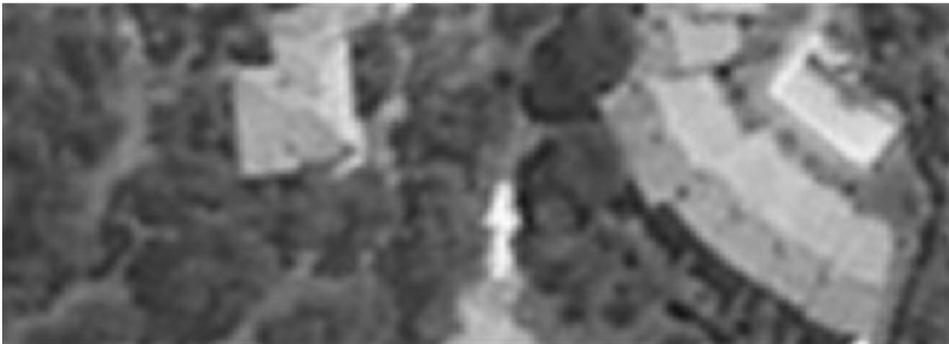


Figura 4. Immagine Google Earth ricampionata.



Figura 5. Contributi puntuali della covarianza.

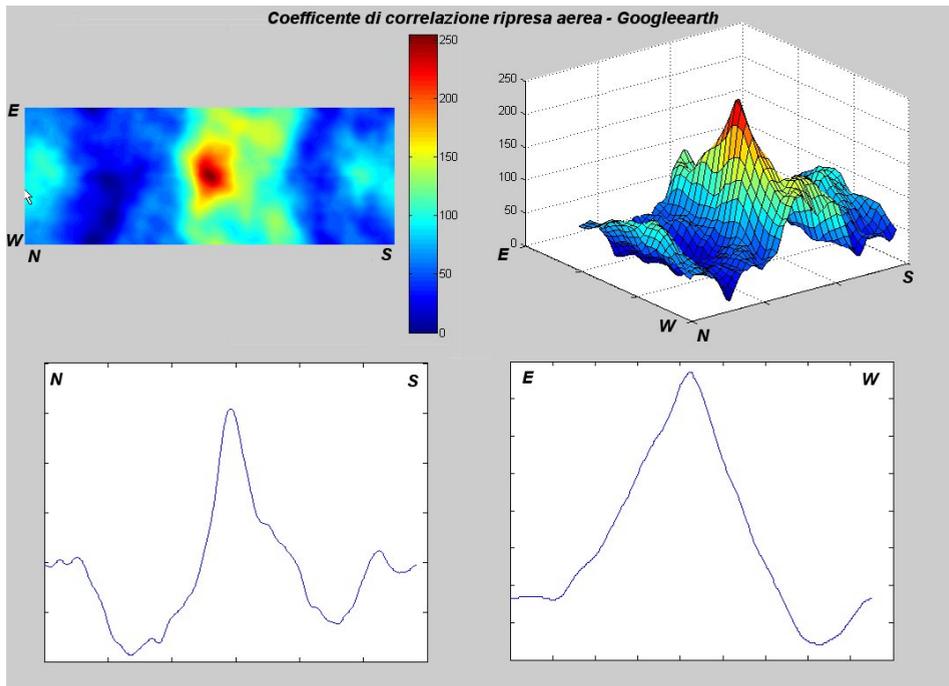


Figura 6. Coefficiente di correlazione tra le immagini di Figura 2 e Figura 4.

Bibliografia

- Bianchi R., Cavalli R.M., Fiumi L., Marino C.M., Pignatti S. (1997), "Airborne remote sensing: Results of two years of imaging spectrometry for the study of environmental problems". Remote sensing 96: integrated applications for risk assessment and disaster prevention for the Mediterranean. Anna Spiteri Editor.
- Capel D. (2004), Image Mosaicing and Super-resolution. Springer-Verlag London.
- Cenedese A., Gabellieri M., Paolillo A., Sciortino R. (2007), "Una metodologia per il monitoraggio ambientale tramite immagini telerilevate iperspettrali e termiche". Atti 11a Conferenza Nazionale ASITA, Torino 6-9 Novembre 2007.
- Gonzales R.C., Woods R.E. (1993), *Digital Image Processing*. Addison-Wesley Publishing Company.
- Goshtasby A. (1986), Piecewise linear mapping functions for image registration. *Pattern Recognition* 19(6): 459-468.
- Lillesand T., Kiefer R.W., Chipman J.W. (2004), *Remote Sensing and Image Interpretation*. Fifth Edition, International Edition, John Wiley & Sons, New York.
- Minacapilli M., Cammalleri C., Ciralo G., D'Asaro F., Iovino M., Maltese A. (2012), Thermal Inertia Modeling for Soil Surface Water Content Estimation: A Laboratory Experiment. *Soil Science Society of America Journal*, 76(1): 92-100.
- Moroni M., Cenedese A. (2012), Analisi di immagini iperspettrali nel monitoraggio ambientale. *L'Acqua* 4-2012.
- Moroni M., Dacquino C., Cenedese A. (2012), Mosaicing of hyperspectral images: the application of a spectrograph imaging device. *Sensors* (doi: 10.3390/s120810228).

- Sims A.D., Gamon J.A. (2002), "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages". *Remote Sensing of Environment* 81: 337-354.
- Stocking M.A., Clark R. (1999), "Soil productivity and erosion: biophysical and farmer-perspective assessment for hillslopes". *Mountain Research and Development* 19(3): 191-202.
- Tapete D., Casagli N., Fanti R., Del Ventisette C., Cecchi R., Petrangeli P. (2012), "Analisi radar satellitare delle deformazioni recenti nell'area archeologica della Domus Aurea e del Colle Oppio (Roma)". *Arkos* 26: 40-49.
- Tatzer P., Wolf M., Panner T. (2005), "Industrial application for inline material sorting using hyperspectral imaging in the NIR range". *Real-Time Imaging* 11(2): 99-107.