

RIVONA Lab : un laboratorio e una FOSS Processing and Archiving Facility per dati iperspettrali

Maria Adamo, Cristina Tarantino, Francesco P. Lovergine, Guido Pasquariello

ISSIA - CNR, Via Amendola 122/D-O, Bari, Italy

Tel +390805929437, fax +390805929460, {cristina, f.lovergine, adamo, pasquariello}@ba.issia.cnr.it

Riassunto

Nel triennio 2010-2012 la Regione Puglia ha finanziato fra le Reti di Laboratori di Ricerca la realizzazione di un laboratorio per la “RIDUZIONE DEI RISCHI DI VOLO E NOWCASTING AEROPORTUALE (RIVONA)”, che prevede l'acquisizione di un sensore iperspettrale aviotrasportato e la realizzazione di un laboratorio scientifico per il processamento dei dati ottenuti da tale sensore. Tale laboratorio vuole caratterizzarsi anche e soprattutto per un uso e sviluppo prevalente di free e open source software per tutte le esigenze di elaborazione e gestione.

Nel corso del primo periodo del progetto, stante l'indisponibilità del sensore e della infrastruttura dovuti ai tempi tecnici di acquisizione degli stessi, il laboratorio ha iniziato comunque una attività di esplorazione delle potenzialità di uso di software FOSS (principalmente, ma non esclusivamente, GRASS), come base per la messa a punto di un workflow di processamento di grandi quantità di dati con elevato numero di bande, utilizzando i dati di un volo iperspettrale del 2009 sulla costa pugliese (complessivamente circa 500GB di dati multibanda). Il workflow proposto è orientato alla realizzazione di un framework automatico di ingestione e pre-processing dei dati, per la costruzione di un catalogo completo di metadati, che risulti facilmente fruibile dagli enti partecipanti al progetto. Come caso test di elaborazione, il framework permetterà di processare la banca dati per la produzione di mappe di SST (*Sea Surface Temperature*) e rendere i risultati disponibili come prodotto applicativo di interesse.

Abstract

In the period 2010-2012 the Apulia Region administration funded the startup of a scientific laboratory named “RIVONA” among other networks of scientific labs. It previsions the acquisition of a hyperspectral airborne sensor and the development of a laboratory facility for its data processing. That laboratory has been qualified even and above all for prevalently using FOSS software in all phases of data processing and management. In the first project period, the laboratory started a study phase to exploit effectiveness of FOSS software (mainly GRASS) to fulfill a processing workflow for a largedatabase of multi-bands data. We used data caught in 2009 for a regional airborne mission (about 500GB of rasters) along the Apulia coast. The workflow is intendend to create an automatic ingestion and pre-processing system, in order to build a complete catalog of metadata, and let it be easily browseable by all partecipants to the project. The framework targets also the extraction of SST (*Sea Surface Temperature*) maps from the whole data set, and making them available along with the other data.

Introduzione

La Regione Puglia ha recentemente finanziato su base triennale (2010-2012) la realizzazione di alcune reti di laboratori di ricerca e in particolare la realizzazione di un laboratorio interdipartimentale del CNR (i tre istituti nazionali coinvolti sono ISSIA, ISAC e IRSA) relativo alla sperimentazione nell'ambito della “Riduzione dei rischi di volo e nowcasting aereoportuale”, in

sigla RIVONA (cfr. Accordo di Programma Quadro in materia di ricerca scientifica tra Ministero dello Sviluppo Economico, Ministero dell'Università e della Ricerca e Regione Puglia: Ricerca Scientifica; "Proposte di piani relativi alle reti di laboratori pubblici di ricerca", progetto 75). Tale laboratorio prevede tra l'altro la realizzazione di una *facility* condivisa per l'archiviazione e il processamento a fini scientifici di dati acquisiti da un sensore iperspettrale aviotrasportato (VNIR 380-1050 nm e TIR 8-12nm). Il laboratorio, fin dalla sua formulazione iniziale, intende caratterizzarsi nel segmento di *processing* anche e soprattutto per un uso e sviluppo prevalente di e con software libero per tutte le esigenze di elaborazione dei dati. Nel corso del primo periodo del progetto, sono state individuate le specifiche infrastrutturali della *Processing and Archiving Facility (PAF)* ed i software da impiegare nel laboratorio per la ingestione e il processamento della grande quantità di immagini che prevedibilmente il PAF dovrà essere in grado di trattare. Dal momento che l'infrastruttura medesima è in fase di acquisizione e messa a punto e lo stesso sensore non è ancora disponibile, è iniziata una fase di sperimentazione preliminare impiegando un corpuso *dataset* di proprietà della Regione Puglia relativo a un volo 2009 del sensore iperspettrale MIVIS (Blom-CGR, 2005) lungo l'intera costa pugliese. L'obiettivo è stato quello di sperimentare e mettere a punto un sistema di processamento il più possibile automatico e da ritenersi "tipico" per mole di dati e caratteristiche dei dati stessi (e quindi connesse problematiche di gestione ed elaborazione) e in qualche modo analogo a quello che il PAF sarà chiamato ad effettuare a regime. Questo ha permesso di prendere dimestichezza con i tool FOSS più adatti e definire soluzioni percorribili sulla base di un *test bed* realistico. I tool utilizzati sono stati prevalentemente il GIS GRASS (Neteler et. al., 2008) e le utility di GDAL/OGR. Alcune procedure sono state messe a punto inizialmente utilizzando il software proprietario ITT ENVI© su frame campione e poi convertite in apposite set di script per il processamento intensivo dell'intero dataset.

Caratteristiche del sensore MIVIS e dati

Il sensore iperspettrale utilizzato è il MIVIS (Multispectral Infrared and Visible Imaging Spectrometer), un sistema a scansione costituito da quattro spettrometri che acquisiscono simultaneamente la radiazione proveniente dalla superficie terrestre nel visibile (20 bande tra 0,441 e 0,819 μm), nel vicino infrarosso (8 bande tra 1,145 e 1,54 μm), nell'infrarosso medio (64 bande tra 1,992 e 2,474 μm) e nell'infrarosso termico (10 bande tra 8,34 e 12,42 μm) per un totale di 102 canali. Dall'applicazione della calibrazione radiometrica a ciascun pixel nelle bande 1 - 92 è associato il valore di radianza espressa in $\mu\text{W}/\text{cm}^2 \cdot \text{sr} \cdot \text{nm}$ mentre per le bande nell'infrarosso termico (canali 93-102) i valori nei pixel sono convertiti in temperatura di brillantezza al sensore espressa in C. Il sensore è dotato di una buona risoluzione radiometrica (12 bits) e spaziale. Essendo infatti un sensore aviotrasportato, nel caso specifico del dataset analizzato ad una quota di 1.5 Km, possiede un pixel di dimensione 3 m x 3 m al nadir. Le acquisizioni MIVIS del dataset sono tutte diurne e sono state effettuate in un periodo temporale che va dal 21 al 25 Maggio 2009.

In corrispondenza dei rilievi iperspettrali da aereo sono state effettuate delle acquisizioni di dati di verità a mare in 4 regioni marine in prossimità dello sbocco a mare di sorgenti. In ogni area sono disponibili 18 campionamenti lungo 6 transetti ortogonali alla costa. In ogni stazione sono stati rilevati profili multiparametrici lungo la colonna. In particolare i dati di temperatura più prossimi alla superficie del mare (circa 0,2 m - 0,3 m di profondità a seconda della stazione) sono stati utilizzati per implementare e validare la procedura di estrazione di mappe di SST.

Estrazione di frame e multilook RGB

L'intero *dataset* - costituito da 97 file in formato ENVI Image BSQ (*Band Sequential Format*) Geocoded multibanda in proiezione Gauss-Boaga Zona 2 con datum Roma 1940 - è stato caricato in un *geodb* GRASS e da tutti i *raster* sono state ricavate maschere e vector *skeleton* sia nella proiezione originale, che in coordinate geografiche su ellissoide WGS84, e in proiezione UTM Zona 33/WGS84. A ciascuno *skeleton* sono stati associati tutti i metadati disponibili, estratti in parte dai file raster originali, in parte dalla documentazione accessoria. Dall'intero set di *skeleton* è stato

ricavato un file KML, consultabile con GoogleEarth, contenente tutti i metadati e gli URL del *repository* interno di archiviazione dei dati originali, per consentire una rapida consultazione da parte di tutti i partecipanti alle attività RIVONA. I raster originali hanno una dimensione variabile fra 500MB e 30GB ciascuno. Sono stati ricavati da alcune bande nel visibile, nel dettaglio i canali 12-7-2 composti rispettivamente come RGB, e per tutti i frame disponibili raster di *multilook* a varie risoluzioni da integrare in un *webgis* di successiva implementazione per la semplice navigazione del dataset.

Estrazione della linea di costa e mask relativo

Primo passo per l'elaborazione dell'immagini MIVIS per l'estrazione delle mappe di SST è stata l'estrazione della linea di costa e la creazione di una maschera terra/mare, in modo da poter facilmente isolare i pixel di mare. L'algoritmo di *land masking* utilizzato è un semplice algoritmo a soglia applicato ai valori di temperatura di brillantezza misurati nella banda 94. Si sfrutta la condizione per cui la temperatura del mare è sensibilmente più bassa di quella della terra. L'algoritmo esegue una prima selezione con una soglia stimata a seguito di opportune prove empiriche e, una volta individuate la regione di mare e quella di terra, ripulisce queste regioni da eventuali zone corrispondenti ad un errato riconoscimento, come ad esempio il caso di piccoli agglomerati di pixel relativi a navi che l'algoritmo di sogliatura etichetta come terra poiché caratterizzate da una maggiore temperatura di brillantezza rispetto al mare circostante. Dall'applicazione di questo primo step di processing si ottiene una immagine binarizzata. Si è quindi ottenuta una linea di costa pugliese aggiornata al 2009 e ad una risoluzione sensibilmente migliore rispetto a quelle normalmente a disposizione con risoluzioni dell'ordine del centinaio di metri.

Estrazione e analisi dei dati termici.

La SST è un importante parametro in molti ambiti di ricerca e per questo motivo la stima di tale parametro attraverso tecniche di telerilevamento ha sempre interessato la comunità scientifica. Molti strumenti utilizzati per questo scopo sono spettrometri operanti nelle finestre atmosferica tra 3,5-4 μm e 8,0-14,0 μm , sia orbitanti intorno alla Terra a bordo di satelliti, sia aviotrasportati. In letteratura sono presenti diversi lavori in cui sono stati sviluppati algoritmi di *retrieval* della SST a partire da immagini satellitari (Brown e Minnett, 1999, O'Carroll et al. 2004) utilizzando diversi tipi di approcci, mentre in (Pepe et al., 2001) sono state ottenute mappe di temperatura sul Lago di Iseo a partire da dati MIVIS utilizzando una regressione lineare multipla con la banda 97 (10 μm) e il suo quadrato. La procedura di estrazione di mappe di SST dalle immagini MIVIS sfrutta le bande nel TIR in cui opera lo spettrometro ad infrarosso di cui è dotato il sensore. Una valutazione della qualità dei dati disponibili è stata preliminarmente condotta con l'obiettivo di riconoscere bande che potessero eventualmente presentare un livello di rumorosità maggiore rispetto alle altre. Questo studio è stato effettuato considerando la deviazione standard di aree omogenee di mare selezionate su alcune immagini del dataset. Dall'analisi effettuata è risultato che le bande 101 e 102 rispettivamente a 11.924 μm e 12.42 μm sono caratterizzate da un rumore sensibilmente maggiore rispetto a quello che caratterizza le restanti bande per cui non sono state utilizzate nell'algoritmo di stima della SST.

Nella regione spettrale del TIR la radianza acquisita dal sensore L_{sat} è funzione della temperatura del corpo reale alla superficie (T) e dell'emissività (ϵ_λ) secondo la relazione:

$$L_{sat} = \epsilon_\lambda \cdot B_\lambda(T) \quad [1]$$

dove B_λ è la radianza che emette un corpo nero avente la temperatura T ad una data lunghezza d'onda e che soddisfa la legge di Plack. Questa è modificata con un termine moltiplicativo dalla trasmissione in atmosfera più il contributo additivo di radianza upwelling.

In questa analisi partiamo dall'assumere che l'atmosfera non incida in maniera significativa sulla radiazione che raggiunge il sensore, considerando il fatto che la quota del sensore è di soli 1.5 Km.

L'informazione contenuta in ciascun pixel dell'immagine calibrata è la T_b , la temperatura che dovrebbe avere un corpo nero per emettere la radiazione L_{sat} . Dalla legge di Planck risulta quindi:

$$\begin{cases} L_{sat} = \frac{C_1}{\lambda^5 \pi \left(e^{C_2/\lambda T_b} - 1 \right)} \\ B_\lambda(T) = \frac{C_1}{\lambda^5 \pi \left(e^{C_2/\lambda T} - 1 \right)} \end{cases} \quad [2]$$

dove C_1 e C_2 sono due costanti rispettivamente pari a $3.7418 \times 10^{-16} Wm^2$ e $1.43877 \times 10^{-2} mK$.

Dalla [2] e [3] consegue che si potrà esprimere la T che nel nostro caso specifico coincide con la SST come:

$$T = \frac{C_2}{\lambda \ln \left[\varepsilon_\lambda \left(e^{C_2/\lambda T_b} - 1 \right) + 1 \right]} \quad [3]$$

Selezionando in maniera random all'interno del dataset di dati di verità a mare un sottoinsieme di training costituito dal 50% della totalità dei dati disponibili si è applicato una procedura ai minimi quadrati, che per abbreviazione chiameremo MLS (MIVIS Least Squares), per determinare, per ciascuna delle 8 bande utilizzate, il miglior valore di ε_λ che posto nella [3] consente la migliore stima di SST. I valori assegnati alla T_b sono quelli relativi ad una finestra 3x3 centrata sul pixel corrispondente alla posizione geografica in cui è stata effettuata la misura.

Dall'analisi svolta sul dataset di training si sono ottenuti i valori di ε_λ riportati in Tab. 1 in cui sono anche riportati i valori teorici ricavati attraverso una modellizzazione della superficie marina in (Masuda et al., 1988) per acqua marina, vento medio pari a 5 m/s, angoli di vista del sensore sino a circa 40° e lunghezze d'onda in un range tra 3,5 μm e 13 μm . I valori di ε_λ relativi alle lunghezze d'onda delle bande del sensore sono stati ottenuti come interpolazione dei precedenti.

ε_λ	Banda 93 (8.340 μm)	Banda 94 (8.748 μm)	Banda 95 (9.179 μm)	Banda 96 (9.571 μm)	Banda 97 (10.000 μm)	Banda 98 (10.420 μm)	Banda 99 (10.933 μm)	Banda 100 (11.428 μm)
MLS	0.96449	0.96670	0.97048	0.97499	0.97637	0.97717	0.97979	0.98028
Masuda et al.	0.98385	0.98487	0.98595	0.98693	0.98900	0.99068	0.99230	0.99207

Tabella 1 – Emissività dell'acqua marina ottenuta in Masuda et al. e dall'applicazione del MLS.

La stima di ε_λ ottenuta applicando MLS risulta essere sempre inferiore a quella stimata teoricamente. Ciò potrebbe essere spiegato ipotizzando un contributo al termine moltiplicativo ε_λ dovuto all'assorbimento atmosferico in termini di trasmittanza. Tale contributo risulta essere compatibile con i valori presumibilmente attesi (Robinson, 2004).

Si è infine utilizzato il restante 50% del dataset come test per validare la procedura di stima della SST e si sono confrontati i risultati, in termini di RMSE (Root Mean Square Error) e di Bias, con quelli che si ottengono utilizzando i valori teorici di ε_λ o utilizzando direttamente la temperatura di brillantezza T_b misurata dal sensore (Fig. 2).

Associare alla temperatura di brillantezza il significato di SST, comporta un errore nella stima non accettabile. Pertanto risulta necessaria una correzione sull'emissività. Utilizzando i valori di emissività ricavati teoricamente si ottengono delle stime di SST con un errore medio di circa 1 C e un Bias negativo per ogni lunghezza d'onda utilizzata che denota una sistematica sottostima della SST. Ciò potrebbe essere interpretato come l'effetto dell'assorbimento dell'atmosfera di cui sino ad ora non si è tenuto conto. L'utilizzo dell' ε_λ ottenuto dall'applicazione di MLS consente di stimare la SST con un RMSE di circa 0,6 C e un bias sensibilmente ridotto che non mostra un trend sistematico di sottostima. L'ipotesi iniziale di trascurare il contributo atmosferico considerando la bassa quota di volo del sensore non è quindi accettabile e la nuova stima di ε_λ consente di

migliorare il *retrieval* della SST perché in essa è come se fosse contenuto il termine di assorbimento. Per ottenere le mappe di SST si utilizzerà il valor medio delle stime ottenute per le 8 bande utilizzate in modo da ridurre il rumore a cui è soggetta la singola banda.

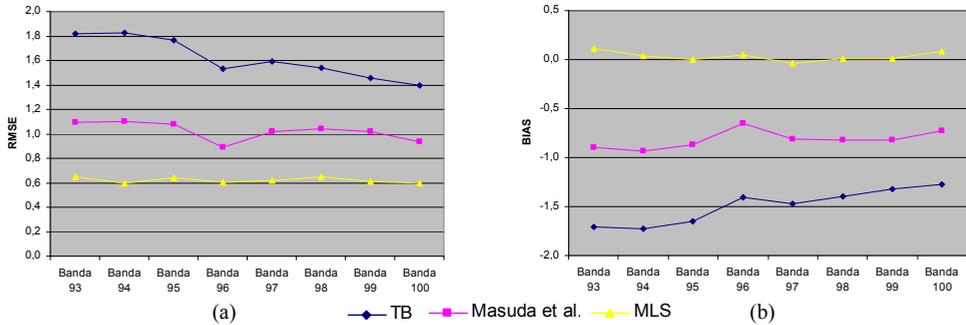


Figura 2 – Andamento della RMSE (a) e del BIAS (b) sulla stima della SST al variare delle bande MIVIS. In blu è rappresentata la curva relativa alla temperatura di brillantezza, in rosa alla SST ottenuta utilizzando i valori di emissività in Masuda et al., in giallo utilizzando le stime di emissività ricavate con MLS.

Il framework di processamento dei dati.

Il sistema di processamento del dataset MIVIS è sinteticamente illustrato nelle sue componenti essenziali in figura 3 e si basa sulla implementazione di un set di *script* per GRASS e l'uso del *tool* di GDAL volti alla estrazione, conversione e analisi dei dati e metadati disponibili; alla produzione delle mappe SST secondo la procedura illustrata in precedenza; alla creazione di un catalogo di vector e raster in multirisoluzione destinato alla visualizzazione e analisi qualitativa di sintesi dei risultati e dei dati di partenza. Parallelamente alle mosaicature in tricromia nello spettro del visibile, è stato prodotto un file in formato OGC KML per una comoda consultazione anche off-line dei dati disponibili e dei risultati di sintesi. L'intero catalogo è destinato all'impiego in un sistema webgis, oggetto di implementazione in una successiva fase del progetto.

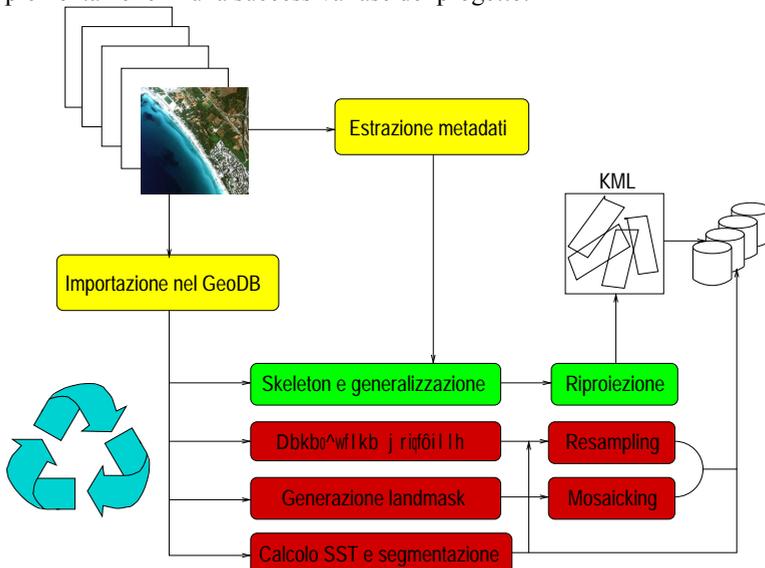


Figura 3 – Schema dei moduli di *processing* per l'applicazione SST e la pre e post elaborazione dei dati.

Sperimentazione.

L'intera procedura è stata applicata al dataset delle immagini MIVIS. Nella Fig. 4 è rappresentata, come esempio, una composizione di prodotti in uscita ai vari step della catena di processing relativi ad una acquisizione del 24 Maggio 2009 alle ore 08:40 locali.

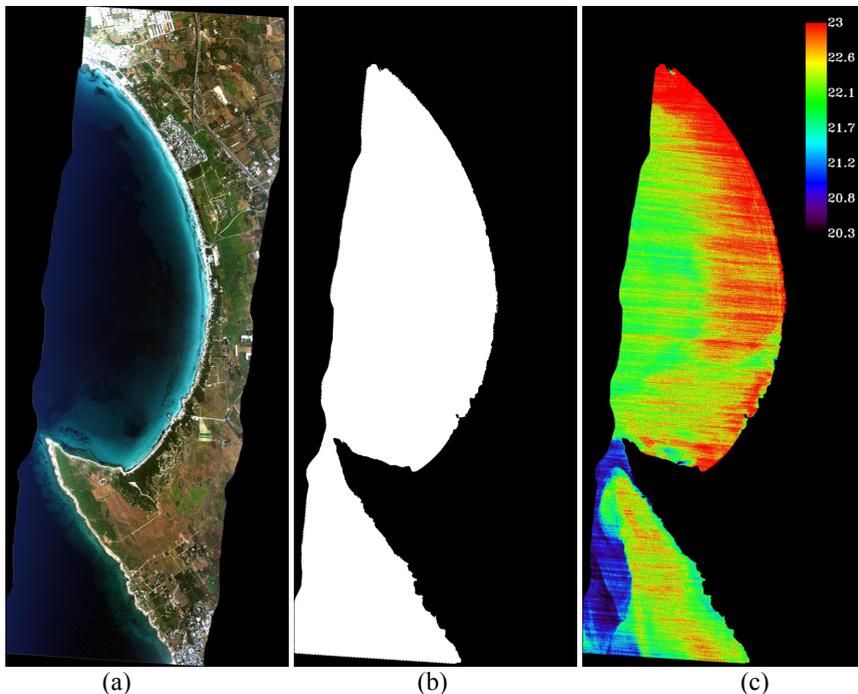


Figura 4 – Immagine MIVIS acquisita 24 Maggio 2009 alle ore 08:40 locali: (a) composizione RGB 12-7-2, (b) maschera di terra, (c) mappa di SST.

Ringraziamenti

Si ringrazia Planetek Italia s.r.l. per aver fornito i dati del sensore iperspettrale MIVIS e quelli di verità in situ ed in particolare la Dott.ssa D. Iasillo e il Dott. G. Ceriola per la loro cortese collaborazione.

Bibliografia

- Blom-CGR (2005), *Scanner Iperspettrale MIVIS*, Url: <http://www.cgri.it/tecnologie/mivis.html>
- Brown, O. B., & Minnett, P. J. (1999). *MODIS infrared sea surface temperature algorithm—algorithm theoretical basis document. Products: MOD28. ATBD-MOD-25.*
- Masuda K., T. Takashima, and Y. Takayama, *Emissivity of pure and sea waters for the model sea surface in the infrared window region*, Remote Sens. Environ., vol. 24, pp. 313–329, 1988.
- Neteler M., Mitasova H. (2008), *Open source GIS : A GRASS GIS Approach*, Ed. Springer.
- O’Carroll, A. G., J. G. Watts, L. A. Horrocks, R. W. Saunders, and N. A. Rayner, *Validation of the AATSR Meteo product sea surface temperature*. *J. Atmos. Oceanic Technol.*, 23:711–726, 2006.
- Pepe M., C. Giardino, E. Zilioli, *Applicazione dei dati MIVIS al monitoraggio dei laghi: discussione su alcuni risultati sperimentali e prospettive*, Rivista Italiana di Telerilevamento, n° 20/21, 2001
- Robinson I.S., *Measuring the Oceans from Space, The principles and methods of satellite oceanography*, Springer, 2004.
- Warmerdam F. et al. (1998), GDAL/OGR, Url: <http://www.gdal.org/>.