

Classificazione di dati telerilevati ad alta risoluzione: risultati di approcci contestuali in ambito Open Source

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

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

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

Abstract

La classificazione di immagini telerilevate ha un'importanza fondamentale per l'estrazione di informazioni utili dai dati. Se si dispone di dati di verità a terra sulla scena è possibile eseguire classificazione di tipo supervisionato: in questo contesto, è possibile utilizzare metodologie *pixel-wise* oppure *object-oriented*. Il primo tipo utilizza le informazioni spettrali associate nell'immagine a ciascun pixel, il secondo tipo utilizza le informazioni spettrali unitamente a descrittori topologici, tessiturali, relazionali o contestuali associati ad un oggetto costituito da un gruppo di pixel omogenei. L'impiego di informazioni contestuali è particolarmente rilevante per immagini ad elevata risoluzione spaziale. La tendenza crescente è quella di rivolgersi al mondo del software libero per attingere a risorse non solo a costi minimi, ma spesso potenzialmente più versatili e sempre aperti. Ne è un esempio GRASS, che fra i suoi moduli mette a disposizione classificatori supervisionati (e non), anche contestuali. Con l'intento di estrarre mappe di uso del suolo da dati telerilevati ad alta risoluzione mediante classificazione supervisionata, in questo lavoro si stima l'efficacia dell'algoritmo di classificazione SMAP di GRASS, che affianca all'informazione spettrale *pixel-wise* anche quella contestuale, confrontandolo con il più convenzionale metodo di classificazione puramente spettrale come il Maximum Likelihood e un classificatore proprietario commerciale. La sperimentazione ha riguardato immagini dei sensori GeoEye-1 e Komsat-2 utilizzati per la produzione di mappe di copertura del suolo all'interno del progetto ASI-MORFEO. Nel lavoro sono discussi i risultati del confronto tra i due approcci.

Classification of remote sensed images has a fundamental relevance to extract useful information from data. Supervised classification can be performed when ground truth for the acquired scene is available. On those regards, it is possible to exploit either pixel-wise or object-oriented approaches: the former uses spectral information for each pixel in the image, the latter uses both spectral information and topological, textural, relational or context descriptors which belong to the same objects, i.e. groups of homogenous pixels.

The use of context-based information has a specific relevance for high resolution images.

Furthermore, current trend is exploiting FOSS programs to gain also in flexibility and openness while minimizing costs. An example is GRASS which has supervised and unsupervised classifiers (even contextual) among its available modules.

In order to extract thematic maps from high resolution images using supervised classification, in this paper we analyze the effectiveness of the hierarchical SMAP algorithm in comparison with a more conventional pixel-wise Maximum Likelihood method and a proprietary commercial classifier.

Experiments have been conducted with Geoeye-1 images used to produce land cover maps within the ASI-MORFEO project. In this paper, results of the comparison between the two approaches are presented.

Introduzione

Nell'ambito dell'analisi di dati telerilevati al fine di estrarre da essi informazione utile per attività tra le più varie in ambito di monitoraggio e gestione ambientale, prevenzione, etc., è indispensabile tener conto della tendenza evolutiva delle più recenti tecnologie: l'affermarsi di sensori ad alta risoluzione spaziale a discapito di quella spettrale impone il ricorso a metodologie, per l'estrazione di mappe tematiche di copertura del suolo mediante classificazione, che integrino informazione aggiuntiva accanto a quella puramente spettrale associata al singolo pixel. Sono disponibili pochi software proprietari che implementano approcci di tipo *object-oriented*, basati sull'individuazione nell'immagine di segmenti spettralmente omogenei, ai quali poter associare una serie di regole *user-driven* di natura spaziale, relazionale o di contesto. Tali sistemi richiedono anche una notevole competenza da parte dell'operatore. In ambito di tool FOSS per la classificazione di immagini il software GRASS 6.4, si presenta particolarmente versatile anche per quanto concerne la classificazione supervisionata, affiancando all'algoritmo statistico *pixel-wise* Maximum Likelihood (ML) un algoritmo SMAP (Sequential Maximum A Posteriori) ad approccio contestuale (Bouman, Shapiro, 1992). Tale algoritmo consente di sfruttare l'informazione spettrale di contesto a scale diverse per migliorare significativamente e automaticamente l'accuratezza della classificazione.

I risultati relativi al confronto tra la metodologia di classificazione ML e quella SMAP sono stati oggetto di studio del presente lavoro, con un'applicazione ad un data set di immagini telerilevate ad alta risoluzione acquisite dal sensore Geoeye-1 nell'ambito del progetto ASI-MORFEO (Morfeo-Project, web site).

Metodologia

Quando si dispone di una verità a terra significativa è possibile applicare algoritmi di classificazione supervisionata per produrre mappe di copertura del suolo da immagini telerilevate.

L'algoritmo standard ML (Maximum Likelihood) è di tipo statistico basato sulla massimizzazione della probabilità a posteriori $P(\omega_i | \mathbf{X})$ che dato un pixel \mathbf{X} (vettore di N componenti pari al numero di bande) esso appartenga ad una assegnata classe ω_i (con i pari al numero delle classi) (Duda, Hart e Strok, 2006). Secondo la formula di Bayes:

$$\arg \max_{\omega_i} P(\omega_i | \mathbf{X}) = \arg \max_{\omega_i} \frac{p(\mathbf{X}|\omega_i)P(\omega_i)}{p(\mathbf{X})} \quad [1]$$

dove $p(\mathbf{X}|\omega_i)$ che rappresenta la probabilità condizionata di trovare un pixel nota una assegnata classe, è ricavata direttamente dai dati di verità a terra assumendo per essi una distribuzione di probabilità gaussiana multivariata e calcolando i relativi parametri statistici di media e matrice di covarianza per ogni classe e, la $P(\omega_i)$, probabilità a priori della singola classe, è in genere posta come equiprobabile tra le classi. Un'evidente limitazione di ML è l'ipotesi di gaussianità nella distribuzione dei dati che nei casi reali tendono a discostarsi da tale andamento, e la mancanza di un qualsiasi criterio addizionale di ottimizzazione (*prior*) su scala globale o locale.

L'algoritmo SMAP (Sequential Maximum A Posteriori) implementato in GRASS (Grass Manuals), accoppia un approccio di *Gaussian mixture model* per massimizzare la separabilità fra le classi supervisionate, ad un metodo di segmentazione gerarchico e multiscala per la classificazione contestuale. Partendo dagli esempi delle varie classi fornite dall'utente, viene calcolato un nuovo set di classi principali, ipotizzando che ogni classe suggerita dall'utente esprima un modello di combinazione lineare di sottoclassi *hidden* con distribuzioni Gaussiane indipendenti. Le sottoclassi sono calcolate secondo un principio di massima sintesi (*Minimum Description Length*) originariamente concepito da (Rissanen, 1983) nel contesto della teoria dei codici (Grunwald, 2005). Il classificatore SMAP costruisce una rappresentazione a piramide multiscala per la mappa di classificazione a partire dalle sottoclassi e dall'immagine multibanda, con l'obiettivo di segmentare la mappa tenendo anche conto della classificazione dei vicini prossimi di ogni pixel e con un peso che dipende dalla scala (Bouman, Shapiro, 1992). L'accoppiamento tra i pixel risulta

sempre meno rilevante man mano che si passa a scale più *coarse*, essendo invece massimo alle scale più *fine*. Globalmente SMAP minimizza la probabilità di misclassificazione tenendo conto dell'accoppiamento interscala (e dunque contestuale) introdotto. Rispetto ad approcci Markoviani simili, SMAP ha il grande vantaggio di non richiedere approcci di ottimizzazione globale per il rilassamento stocastico, guadagnando in velocità di convergenza.

Data set

Il data set utilizzato è relativo ad una scena acquisita dal satellite Geosy-1 in orbita da settembre 2008, dotato, nel multispettrale, di 4 bande nel visibile e vicino infrarosso, alla risoluzione spaziale di 1.65 metri nominale e di 2 metri effettiva. Le immagini, orto rettificata, utilizzando gli RPC e un DTM a 20 metri, e proiettate nel sistema WGS84-UTM32, sono relative ad un sito tra i comuni di Varenna ed Esino Lario in Lombardia (Italia) test site nel progetto ASI-MORFEO.



Figura 1 – Composizione RGB del test site utilizzato: un'area tra i comuni di Varenna ed Esino Lario in Lombardia.

Sperimentazione e Risultati

La sperimentazione è stata basata sul confronto tra la metodologia di classificazione puramente spettrale ML, presente nel software per *image processing* ENVI, e quella spettrale e contestuale SMAP, presente nel software GRASS/FOSS, utilizzando come figure di merito il valore di Overall Accuracy (OA%) e del coefficiente Kappa (K) (Cohen, 1960), (Congalton e Green, 1999).

La verità a terra è stata acquisita direttamente per fotointerpretazione di ortofoto della regione Lombardia disponibili sul web. Sono state individuate nella scena 6 classi principali: Latifoglie – Arborato – Pascolo – Nudo – Acqua – Urbano ed acquisiti per ciascuna di esse un certo numero di poligoni per un totale di circa 20000 pixel (training set - TR). Unitamente è stato individuato sulla scena un data set di test, composto da poligoni relativi alle 6 classi per un totale di oltre 130000 pixel (test set – TE) ben distribuiti sull'intera scena, sul quale valutare gli indicatori di performance dei classificatori. Sono state eseguite le seguenti classificazioni fornendo in input le 4 bande nelle frequenze del multi spettrale:

1. Classificazione con ML, a 6 classi, disponibile in ITT ENVI©
2. Classificazione con ML, a 6 classi, disponibile in GRASS

Tale primo confronto ha consentito di accertare la piena compatibilità tra i due diversi software che hanno fornito esiti assolutamente simili a meno di una diversità di approssimazione dell'ordine dello 0.01% in termini di OA (Tabella 2, casi: A e B);

3. Classificazione con ML a 41 classi accorpate a 6 dopo la classificazione

Per ciascuna classe, sulla base della dispersione delle statistiche attribuite, si è individuato un insieme di sottoclassi $\{\omega_k\}$ rappresentative della stessa classe ω e per ciascuna sottoclasse è stata calcolata la *feature* statistica rappresentativa. In tal modo sono state individuate un totale di 41 sottoclassi e si è prodotta la mappa classificata relativa, si è poi proceduto ad accorpate le sottoclassi riferendole alla classe di riferimento in modo da riottenere una mappa a 6 classi. La specializzazione del training set ha comportato un miglioramento di circa il 6% di OA in training. Tuttavia l'eccessiva specializzazione ha comportato una ridotta capacità di generalizzazione giacché restano inevitabilmente "inseguite" le variazioni dovute al rumore (il miglioramento in test è quasi trascurabile: intorno all'1% di OA e inalterato per K) (Tabella 2, caso C).

4. Classificazione con ML, a 6 classi, con l'attribuzione dei pixel isolati mediante criterio del *Majority Vote* applicato su finestre 3x35. Classificazione con ML a 41 classi accorpate a 6 dopo la classificazione con l'attribuzione dei pixel isolati mediante criterio del *Majority Vote* applicato su finestre 3x3

L'applicazione di un criterio di attribuzione a valle della classificazione consente di introdurre informazione contestuale poiché resta attribuita la probabilità predominante nell'intorno: ciò ha comportato un netto miglioramento di circa il 7% di OA e 0.1 di K nel caso a 6 classi (miglior esito tra le prove eseguite con ML) e del 5% di OA e 0.06 di K nel caso a 41 classi poi accorpate (Tabella 2, casi: D ed E). I due casi tra loro forniscono risultati confrontabili. Il miglioramento conseguito con l'utilizzo del *Majority Vote* è indicativo del miglioramento che l'adozione di un algoritmo di tipo contestuale comporterebbe a prescindere dal numero delle classi considerate;

6. Classificazione con SMAP a 6 classi

Tale caso che utilizza sia informazione contestuale che un modello gaussiano misto per le classi individuando le sottoclassi per ciascuna di esse in maniera automatica ha conseguito i migliori risultati sia in training che in test con un miglioramento del 4% di OA e 0.06 di K rispetto al migliore esito con ML (Tabella 2, caso: F);

7. Classificazione con SMAP a 41 classi accorpate a 6 dopo la classificazione

Come già previsto un'eccessiva specializzazione, sia manuale che automatica, del training set comporta ridotta capacità di generalizzazione con esiti in test, sia di OA che di K, che sono inferiori a quelli descritti nel caso precedente (Tabella 2, caso: G).

CASO	6 CLASSI: LATIFOGGLIE – ARBORATO - PASCOLO – NUDO – ACQUA – URBANO 41 CLASSI: VARIE SOTTOCLASSI PER OGNI CLASSE	TR (19504 punti)		TE (130511 punti)	
		OA%	K	OA%	K
A	ML di ENVI a 6 CLASSI	79.44	0.69	70.40	0.52
B	ML di GRASS a 6 CLASSI	79.43	0.69	70.40	0.52
C	ML a 41 CLASSI accorpate a 6 POST CLASSIFICAZIONE	85.58	0.77	71.89	0.52
D	ML a 6 CLASSI + MAJORITY VOTE	86.07	0.78	77.36	0.62
E	ML a 41 CLASSI accorpate a 6 POST CLASSIFICAZIONE + MAJORITY VOTE	92.43	0.87	76.44	0.58
F	ML CONTESTUALE a 6 CLASSI	90.50	0.85	81.56	0.68
G	ML CONTESTUALE a 41 CLASSI accorpate a 6 POST CLASSIFICAZIONE	97.71	0.96	79.28	0.63

Tabella 2 – Esiti classificazione (OA% e K) in training e test conseguiti nelle diverse sperimentazioni.

Analizzando gli esiti conseguiti classe per classe riportati nella matrice di confusione emerge che mentre il riconoscimento di classi spettralmente più omogenee risulta avvantaggiato dall'utilizzo dell'informazione contestuale (Latifoglie – Acqua) al contrario per quelle spettralmente più eterogenee ne risulta penalizzato (Urbano – Pascolo – Arborato): infatti la classe Nudo è confusa con l'Urbano maggiormente dall'algorithm SMAP (Tabella 3) di quanto non avvenga con il migliore caso (D) analizzato con ML (Tabella 4).

TEST	Ground Truth (%)							
	Class	LAT	ARB	PASC	NUDO	ACQUA	URB	Total
Unclassif	0	0	0	0	0	0	0	0
LAT	97.09	40.76	7.40	0.10	0	0	0	65.93
ARB	1.88	38.04	0.12	0	0	0	0	10.00
PASC	0.72	17.60	92.27	0	0	0	0	13.81
NUDO	0.31	3.45	0.14	34.51	0.20	8.66	1.95	
ACQUA	0	0	0	0	99.14	0.12	3.10	
URB	0	0.15	0.06	65.39	0.66	91.22	5.21	
Total	100	100	100	100	100	100	100	

Tabella 3 –Matrice di Confusione in Test per l'algorithm SMAP (Tabella 2, caso:F).

TEST	Ground Truth (%)							
	Class	LAT	ARB	PASC	NUDO	ACQUA	URB	Total
Unclassif	0	0	0	0	0	0	0	0
LAT	89.35	35.65	2.08	0.14	0	0	0	59.76
ARB	8.11	37.21	0.08	0	0	0	0	13.37
PASC	2.22	25.28	97.82	0.10	0	0	0	17.03
NUDO	0.32	1.86	0.02	59.08	0.57	17.98	2.40	
ACQUA	0	0	0	0	95.82	0	2.99	
URB	0	0	0	40.68	3.61	82.02	4.45	
Total	100	100	100	100	100	100	100	

Tabella 4 –Matrice di Confusione in Test per l'algorithm ML a 6 classi + Majority Vote (Tabella 2, caso:D).

In figura 5c è possibile apprezzare la migliore individuazione del fiume e delle latifoglie da parte dell'algorithm SMAP.



Figura 5a –RGB dettaglio scena Geoye-1.

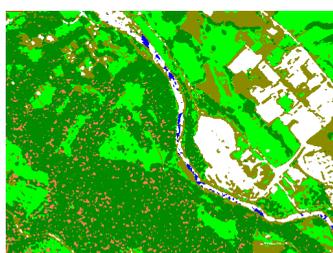


Figura 5b –Classificato ML a 6 classi + Majority Vote.



Figura 5c –Classificato SMAP.

In figura 6b è possibile apprezzare la migliore individuazione della zona di nudo da parte dell'algorithm ML a 6 classi + Majority Vote e la confusione con urbano dell'SMAP.

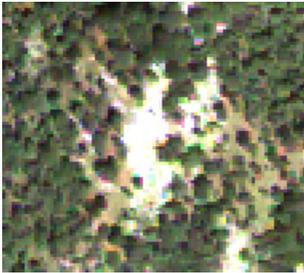


Figura 6a –RGB dettaglio scena Geoeye-1.



Figura 6b –Classificato ML a 6 classi + Majority Vote.

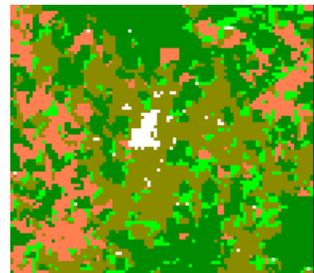


Figura 6c –Classificato SMAP.

Conclusioni

Si è voluta stimare, adoperando tool disponibili *open source*, la differenza nell'utilizzo di due distinte metodologie di classificazione di dati tele rilevati ad alta risoluzione per l'estrazione di mappe di copertura del suolo: spettrale vs. spettrale + contestuale. In particolare l'algorithm spettrale + contestuale utilizzato assume un modello per la distribuzione delle classi di tipo gaussiano misto individuando in maniera automatica una serie di sottoclassi per ogni classe e in tal modo specializzando il data set di verità a terra fornito al classificatore. Tale caso, pur con i limiti presenti nel riconoscimento di classi spettralmente eterogenee, consegue i migliori risultati in test in termini di OA% e di K con una notevole riduzione di spot isolati (misclassificati) nella mappa classificata prodotta che risulta decisamente più uniforme. I tempi di calcolo sono di poco superiori a quelli di un classificatore puramente spettrale.

Bibliografia

- Bouman C., Shapiro M. (1994), "A Multiscale Random Field Model for Bayesian Image Segmentation", *IEEE Trans. on Image Processing*, 3(2): 162-177
- Cohen J. (1960), "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement*, 20(1): 37-46
- Congalton R., K. Green. 1999, "Assessing the Accuracy of Remotely Sensed Data: Principles and Practices", ed. CRC/Lewis Press, Boca Raton
- Duda R.O., Hart P.E. e Strok D.G. (2006), "Pattern classification", ed. Wiley & Sons.
- Grass Manuals, http://grass.itc.it/grass62/manuals/html62_user/i.smap.html
- Grass web site, <http://grass.itc.it/>
- Grunwald P. (2005), "Advances in minimum description length: theory and applications", ed. MIT Press
- Morfeo-Project web site, <http://www.morfeoproject.it/>
- Rissanen J. (1983), "A Universal Prior for Integers and Estimation by Minimum Description Length", *Annals of Statistics*, 11 (2): 417-431