

# Algoritmi di Classificazione “Contextual” con GRASS

Francesco Pirotti, Antonio Vettore, Alberto Guarnieri

\*CIRGEO - Centro Interdipartimentale di Cartografia, Fotogrammetria, Telerilevamento e S.I.T. - Università degli Studi di Padova – Viale dell’Università 16 – 35020 Legnaro (PD) – Tel. 049-8272688/2522 Fax. 049-8272686  
e-mail: cirgeo@unipd.it – <http://www.cirgeo.unipd.it>

## Sommario:

Nell’analisi delle immagini digitali il procedimento di classificazione riveste un’importanza fondamentale per l’estrazione di informazioni utili dai dati grezzi delle immagini. I vari metodi che vengono utilizzati possono essere divisi in due categorie: metodi puntuali e metodi contestuali. In letteratura vengono chiamati anche con altri nomi, ad esempio rispettivamente pixel-based ed object-based. Il primo tipo utilizza solo le informazioni di riflettanza, contenute nell’immagine sotto forma di valori di grigio del pixel classificato, il secondo tipo utilizza anche un altro genere di informazioni (relazioni geo-spaziali, tessitura, contesto, rapporti dimensionali etc...). Questo ultimo genere di classificazione è meno utilizzata per diversi motivi: richiede tempi di preparazione e di elaborazione più lunghi, è a maggior rischio di errore ed infine richiede una preparazione, da parte dell’operatore, più specifica. Ultimamente lo sviluppo di processori a velocità più elevata ha permesso l’implementazione di questo genere di algoritmi in diversi pacchetti commerciali di elaborazione delle immagini digitali telerilevate.

L’obiettivo della seguente nota è di paragonare con GRASS un algoritmo di classificazione che utilizzi il metodo contestuale valutandone l’efficacia rispetto al più convenzionale metodo di classificazione come Maximum Likelihood. Verrà illustrato il metodo di classificazione nelle sue parti, comparandolo ai metodi classici. I risultati verranno comparati tra di loro per vedere quale sia migliore nel contesto del lavoro.

## Introduzione

Per metodi di classificazione convenzionali si intendono algoritmi che utilizzano unicamente l'informazione spettrale dell'immagine. Il problema con questo genere di classificazione è che vengono tralasciate molte informazioni di tipo spaziale e spettrale nell'intorno del pixel classificato. Ad esempio, non vengono considerati i rapporti geo-spaziali all'interno dell'immagine, e neanche il valore spettrale dei pixel intorno a quello considerato. L'interpretazione visiva fatta dall'uomo utilizza molto questo secondo genere di informazioni, ed è per questo che è più efficace rispetto all'interpretazione fatta con calcolatori elettronici che utilizzano solo le informazioni spettrali puntuali. Gli algoritmi tuttora a disposizione per la classificazione di immagini digitali sono per la maggior parte di tipo convenzionale (Cluster Analysis, Maximum Likelihood, Spectral Angle Mapper etc...); gli altri tipi di algoritmi sono ancora poco utilizzati.

I classificatori che utilizzano altre informazioni oltre quelle spettrali puntuali possono essere divisi in due classi: contestuali (contextual classifiers) e basati sulle caratteristiche dell'oggetto (object-based classifiers) in contrapposizione con i classificatori convenzionali anche chiamati a pixel (pixel-based). Questi classificatori utilizzano algoritmi che richiedono un tempo di processore (processor-time, ossia tempo di elaborazione a parità di velocità) decisamente maggiore di quelli convenzionali, in quanto non solo usano la distribuzione dei valori spettrali ma anche le diverse altre informazioni aumentando quindi esponenzialmente la quantità di dati da elaborare. Lo sviluppo della velocità dei processori (e anche l'avvento dei processori a 64 bit) e quindi delle velocità di calcolo ha permesso a questi metodi di classificazione, che prima richiedevano tempi di calcolo spesso proibitivi, di essere implementabili.

I classificatori puramente object-based aggiungono ai metodi classici (region growing, clustering ...) ulteriori criteri di assegnazione alle classi mediante elaborazioni successive. Prima dividono l'immagine in classi solitamente utilizzando il metodo Maximum Likelihood, successivamente le regioni classificate vengono elaborate ulteriormente utilizzando diverse caratteristiche come ad esempio la scala media degli oggetti da estrarre, il grado di prossimità tra classi, la probabilità di vicinanza e gli indicatori di forma e di texture.

I classificatori puramente contestuali invece considerano le caratteristiche spettrali non solo del pixel da classificare, ma anche dei pixel vicini. Per semplicità computazionale vengono utilizzate finestre quadrate con lati di dimensione  $n$  pixels dove  $n$  è un numero dispari. La caratteristica del segnale creato dai valori di grigio all'interno di questa finestra viene rappresentata con diversi indicatori che variano da semplici trasformazioni statistiche (vedi tabella 1) fino a valori

indicanti un aspetto tessiturale (*co-occurrence matrix*) che diversificano le classi in modo da renderle distinguibili dal procedimento di assegnazione.

| Nome  | Espressione Matematica   |
|---|--|
| Media   | $(\sum x_{ij})/n$  |
| Deviazione standard   | $\left\{ \left[ \sum (x_{ij} - AVE)^2 \right] / n \right\}^{0.5}$    |
| Eccentricità  | $\left[ \sum (x_{ij} - AVE)^3 \right] / \left[ n(STD)^3 \right]$     |
| Curtosi   | $\left[ \sum (x_{ij} - AVE)^4 \right] / \left[ n(STD)^4 \right] - 3$ |
| Variazione assoluta dalla media                             | $\left[ \sum  x_{ij} - AVE  \right] / n$                             |
| Indice di contrasto   | $\left[ \sum  x_{ij} - x_c  \right] / (n - 1)$                       |
| Differenza media tra pixel centrale e vicini                | $\left[ \sum (x_{ij} - x_c) \right] / (n - 1)$                       |
| Contrasto tra i vicini                                      | $\left[ \sum  x_i - x_{i+}  \right] / n_+$                           |
| Somma al quadrato del contrasto tra vicini                  | $\left[ \sum (x_i - x_{i+})^2 \right] / n_+$                         |
| Contrasto tra vicini di secondo grado                       | $\left[ \sum  x_i - x_{i++}  \right] / n_{++}$                       |
| Somma al quadrato del contrasto dei vicini di secondo grado | $\left[ \sum (x_i - x_{i++})^2 \right] / n_{++}$                     |
| Range   | <b><i>Max - Min</i></b>  |
| <b>AVE e STD</b>  | Media e Deviazione Standard  |
| $x_{ij}$  | Valore del pixel in posizione ij                                     |
| $x_c$   | Valore pixel centrale  |
| $x_i - x_{i+}$  | Differenza di valore di due pixel vicini                             |
| $x_i - x_{i++}$   | Differenza valore tra due pixel vicini di 2° grado                   |
| $n$   | Numero di pixels nella finestra                                      |
| $N_+$   | Numero di paio di vicini   |
| $N_{++}$  | Numero di paia di vicini di 2° grado                                 |

**Tabella 1 – Indicatori contestuali – Fonte: Prof. Peng Gong - Corso *Advanced Remote Sensing* – UC Berkeley**

Esistono ovviamente procedimenti che fanno uso di metodologie ibride che considerano sia il contesto spettrale del pixel sia le caratteristiche geo-spaziali e le relazioni topologiche.

GRASS ha un modulo di classificazione contestuale *i.smap* che classifica l'immagine utilizzando la Sequential Maximum A Posteriori ossia approfittando il fatto che in un'immagine pixel vicini tendono ad avere un probabilità più alta di appartenere alla stessa classe rispetto ai pixel lontani. Matematicamente questa operazione ipotizza che la distribuzione dei valori nell'immagine

possa essere rappresentata con una distribuzione mista Gaussiana, cioè che contiene al suo interno diverse distribuzioni multivarianza per ogni classe ipotetica. Questo permette di utilizzare sia semplicemente valori di media e varianza per caratterizzare il singolo pixel sia di considerare invece le caratteristiche insieme di tutta un'area.

Nel lavoro presentato verrà valutato l'algoritmo di classificazione Maximum Likelihood rispetto a quello SMAP, ed inoltre verrà proposto un procedimento alternativo che ottimizzi le risorse di GRASS allo scopo di eseguire una classificazione più accurata. Facendo quindi uso di criteri sia classici che contestuali, e verrà fatta una valutazione d'insieme dei singoli procedimenti e di quello proposto. Questi procedimenti verranno applicati per classificare tre diverse immagini: da ETM Landsat, da SPOT, da ASTER.

## Materiali e Metodi

### *Materiali*

In questo lavoro vengono analizzate tre immagini utilizzando i procedimenti proposti: due immagini provengono dal lavoro di Bauman e Shapiro e rappresentano un'immagine da sensore satellitare SPOT ed un'immagine artificiale creata appositamente con caratteristiche particolari; la terza immagine è un'estratto da un'immagine da sensore satellitare ASTER nel veneto. Tra queste tre immagini due sono di immagini satellitari (SPOT e ASTER) mentre l'altra è un'immagine creata artificialmente (vedi figura 1).

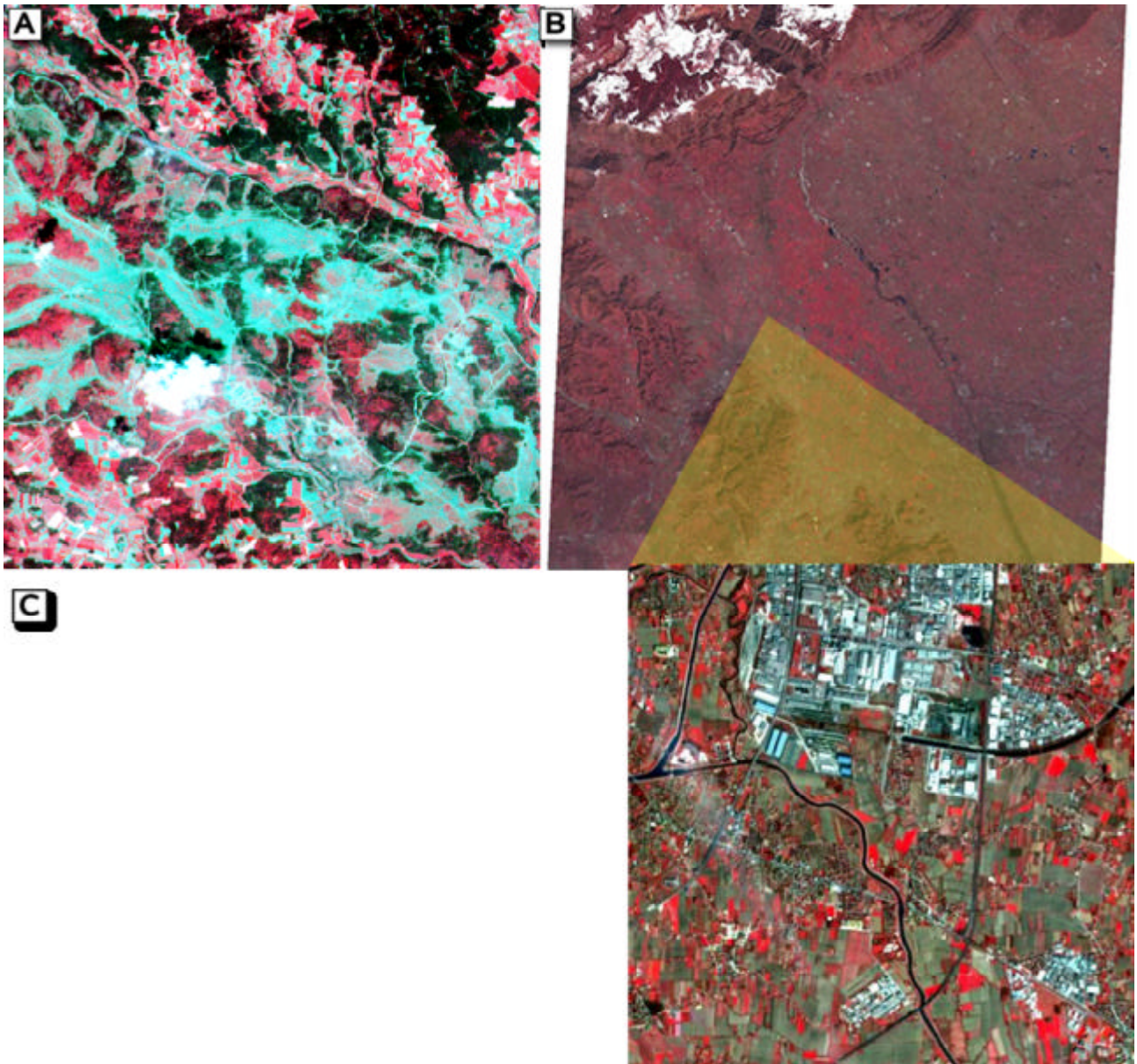


Figura 1 – A: Immagine SPOT. B: Immagine TERRA-ASTER C: Immagine Landsat-ETM+

Queste tre immagini sono state scelte perché presentano caratteristiche peculiari ben diverse tra loro (vedi tabella 2), e quindi rappresentanti di diverse casistiche che un operatore può incontrare.

| Caratteristica | SPOT                                      | ASTER  | LANDSAT   |
|----------------|---|--|---|
| N° BANDE       | 4   | 14   | 3   |
| BIT            | 8   | 8-16   | 8   |
| N° CLASSI      | 5   | 5  | 3   |
| TIPOLOGIA      | Terreno con elementi con forme irregolari | Terreno con strutture sia regolari (edifici, campi ect...) che irregolari (zona) | Immagine con molto rumore e classi spettralmente meno definite. |

**Tabella 2 – Caratteristiche delle tre immagini utilizzate**

Il sensore SPOT utilizza 4 bande per analizzare il territorio usando una zona dello spettro che spazia dal visibile all'infrarosso. Nella nostra immagine, fornita da Bouman e Shapiro, ci sono delle zone individuabili visivamente, ma più che altro si utilizzeranno le aree campione fornite dagli stessi autori per testare il loro algoritmo e paragonarlo ad altri.

Il sensore ASTER è stato montato sulla piattaforma orbitante Terra (EOS AM-1) lanciata nel 1999 e presenta caratteristiche di alto interesse per la classificazione dato che mette a disposizione 14 bande che spaziano dal visibile fino all'infrarosso termico. Ogni immagine copre circa 60 x 60 km ed ha un tempo di ritorno nella stessa area mediamente di 16 giorni. Nel nostro caso è stata presa una parte di immagine di dimensioni 921 pixel x 920 pixel. La zona è rappresentativa di cinque classi ben definite ed è stata scelta per la facile interpretabilità visiva delle classi. Mentre per l'immagine SPOT le aree campione erano già fornite, nel caso dell'immagine ASTER si è proceduto a raccogliere zone campione mediante interpretazione visiva. Le zone da classificare presentano chiaramente diverse classi di notevole interesse per la classificazione della copertura del suolo: strade, territorio urbano, reti idriche, zone agricole, industriali e di montagna. Per la sua risoluzione spaziale, questa immagine presenta alcuni elementi nettamente visibili (campi coltivati, zone urbane) ed altri con esigua presenza di pixel (canali larghi dai 5 – 20 metri, strade) che vengono individuati da interpretazione visiva molto più per la loro forma regolare che per la percezione del colore.

Il satellite Landsat con il sensore Thematic Mapper permette di ottenere una risoluzione di 30 mt nel visibile. Le classi sono così meno definite nei loro contorni a seconda della tipologia di classe. Come nell'esempio dell'ASTER, corsi d'acqua di larghezza superiore ai 30 mt possono essere visualizzati, al di sotto si perde di definizione.

## Metodi

Le immagini sono state classificate utilizzando due metodologie di classificazione: la Maximum Likelihood e la SMAP – Sequential Maximum a Posteriori. Per ogni immagine l'accuratezza dei risultati viene valutata utilizzando il coefficiente Kappa calcolato dalla matrice di confusione. Nel caso di una classificazione perfetta il coefficiente kappa assume valore 1. Esso viene calcolato moltiplicando il numero totale di pixels in tutte le classi (N) per la somma delle diagonalmente della matrice di confusione ( $X_{kk}$ ), sottraendo la somma di tutti i pixels appartenenti alla classe moltiplicati per i pixels associati alla classe considerata (commission) ( $X_k \sum X_k$ ), e dividendo per il numero totale di pixels al quadrato meno lo stesso termine di prima.

$$K = \frac{N \sum X_{kk} - \sum X_k \sum X_k}{N^2 - \sum X_k \sum X_k} \quad (1)$$

Questo termine rappresenta l'accuratezza in modo migliore rispetto alla semplice valutazione della percentuale di pixel correttamente classificati; tiene infatti conto sia degli errori di commissione che di omissione.

La maximum likelihood è un classificatore guidato che utilizza la distribuzione statistica dei valori di grigio delle aree designate per ogni classe per assegnare un grado di probabilità di appartenenza per ogni pixel dell'immagine. La classe del quale il pixel avrà una probabilità maggiore verrà assegnata a quel pixel. I risultati di questo metodo danno risultati soddisfacenti dove le varie classi siano spettralmente ben definite e dove tutti i pixel dell'immagine siano nettamente appartenenti all'una o all'altra classe. Questo non è sempre il caso, anzi, la maggior parte delle volte non lo è.

La SMAP invece utilizza la probabilità di appartenenza ad una classe tenendo conto del risultato della classificazione fatta a priori. Quindi non ci si basa unicamente sulla firma spettrale, ma anche sulla sua vicinanza a pixel classificati in un certo modo. Viene così assegnata una probabilità di appartenenza che è il risultato di considerazioni "contestuali" e spettrali.

L'immagine ASTER viene divisa mediante interpretazione visiva in cinque classi come visibile in tabella 3.

| Classi di informazione | Punti per ogni classe nelle aree campione |
|------------------------|---|
| Area edificata         | 175                                       |
| Prato                  | 161                                       |
| Seminativi             | 151                                       |
| Suolo nudo             | 250                                       |
| Acqua                  | 129                                       |

Tabella 3 – Classi dell'immagine e pixel delle aree campione.

## Risultati ed Analisi

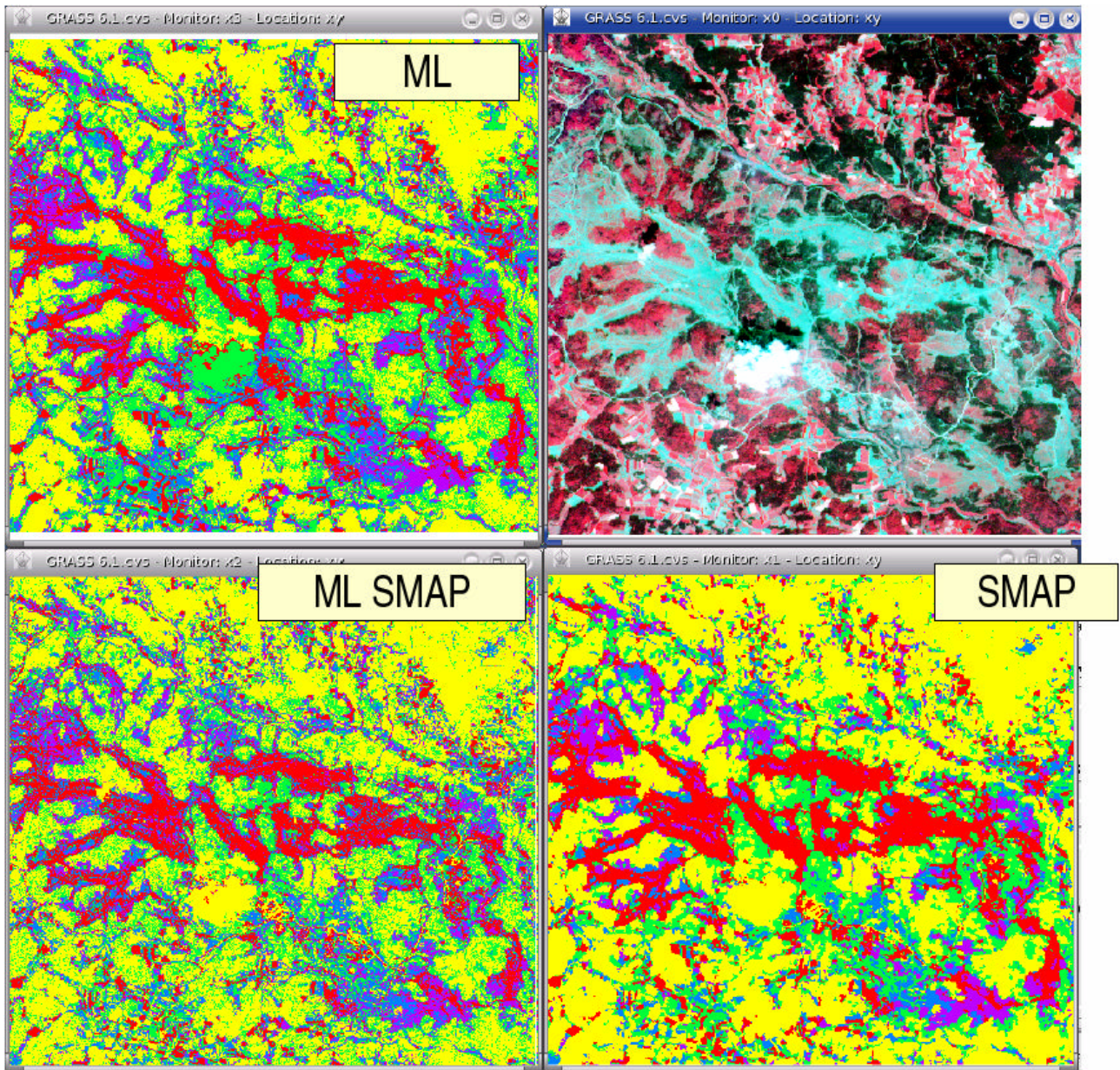


Figura 2 – Risultati con ML e SMAP su SPOT



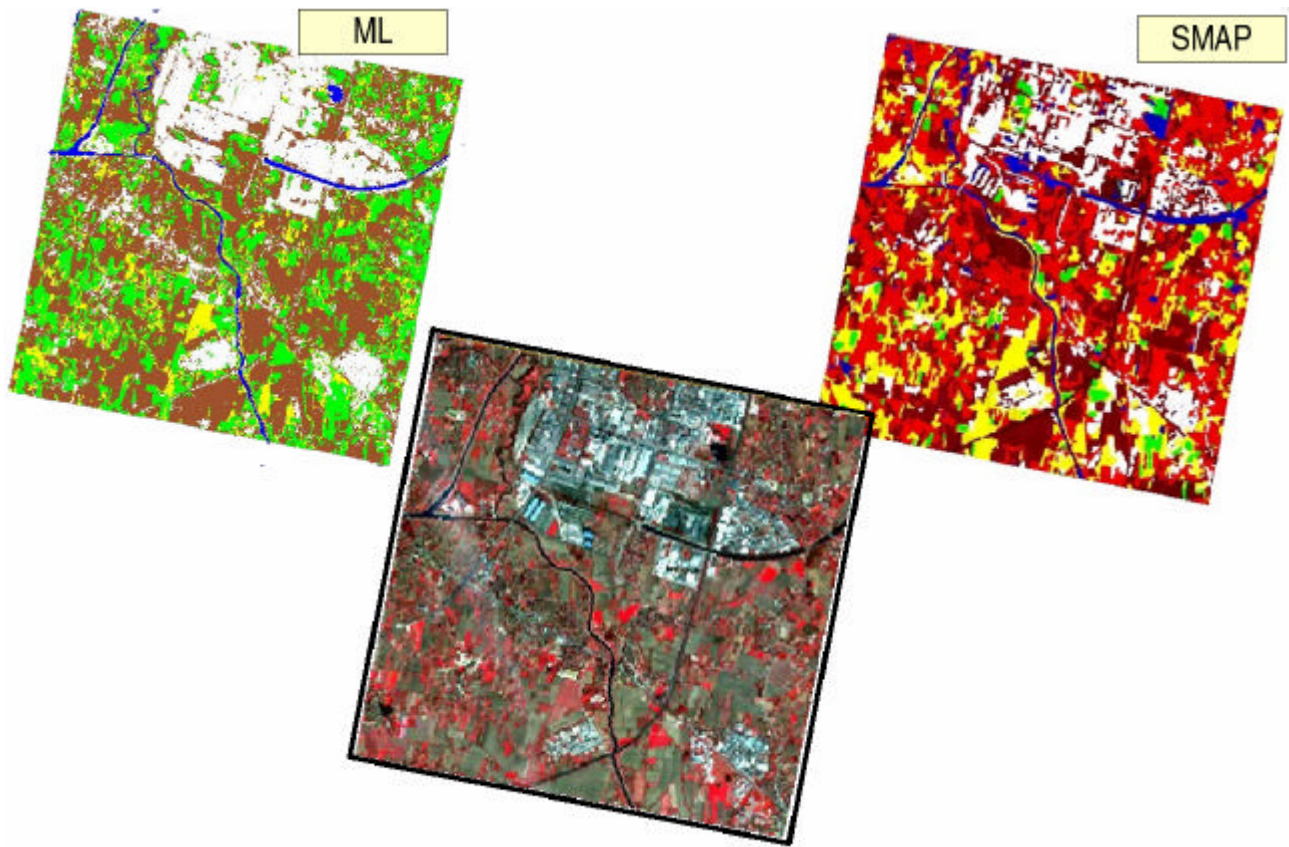


Figura 3 – Risultati con ML e SMAP su ASTER

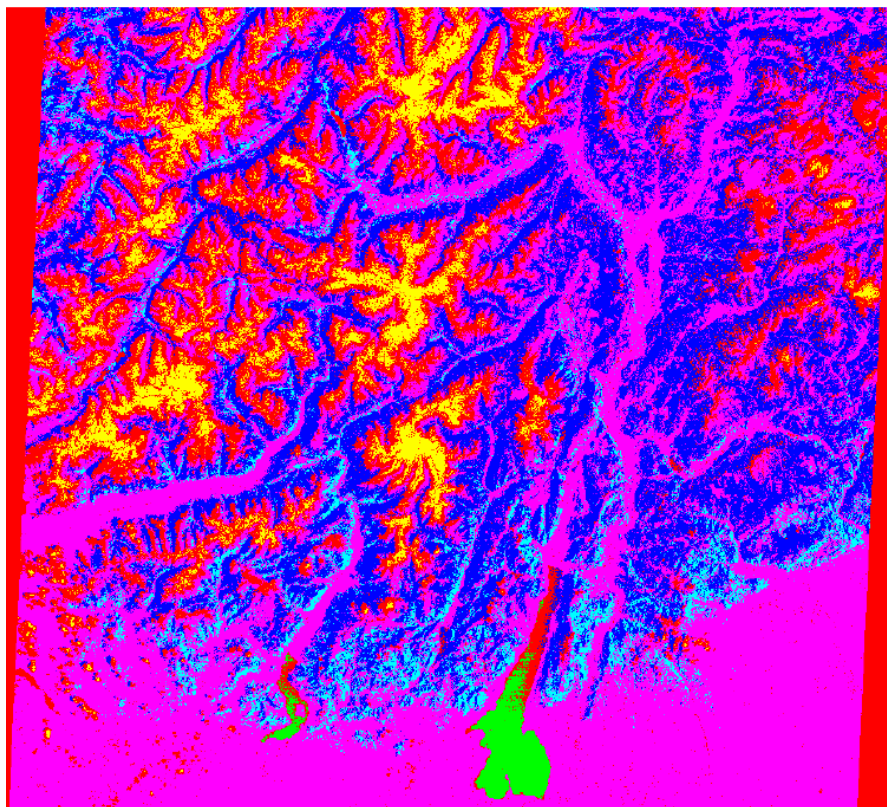


Figura 4 – Risultati con ML su Landsat

| KAPPA Maximum Likelihood |       |         |       |
|--------------------------|-------|---------|-------|
|                          | SPOT  | LANDSAT | ASTER |
| Urbano                   | 0,76  | 0,73    | 0,52  |
| Idro                     | 0,87  | 0,8     | 0,59  |
| Vegetazione              | 0,82  | 0,9     | 0,57  |
| Seminativo               | 0,65  | 0,5     | 0,26  |
| Suolo nudo               | 0,75  | 0,63    | 0,29  |
| media                    | 0,770 | 0,712   | 0,446 |
| dev. standard            | 0,083 | 0,154   | 0,159 |

| KAPPA – SMAP  |       |         |       |
|---------------|-------|---------|-------|
|               | SPOT  | LANDSAT | ASTER |
| Urbano        | 0,82  | 0,9     | 0,64  |
| Idro          | 0,89  | 0,93    | 0,82  |
| Vegetazione   | 0,83  | 0,87    | 0,75  |
| Seminativo    | 0,71  | 0,58    | 0,36  |
| Suolo nudo    | 0,77  | 0,69    | 0,49  |
| medio         | 0,804 | 0,794   | 0,612 |
| dev. standard | 0,068 | 0,152   | 0,188 |

**Tabella 4 - Coefficiente Kappa per le diverse classi utilizzando ML e SMAP**

E' stato interessante mettere a confronto il comportamento dei due classificatori per vedere quale risulti più accurato e specialmente se quest'eventuale miglioramento avviene anche per altre immagini oppure se è correlato a qualche caratteristica dell'immagine. L'indagine è ben lungi dal essere esaustiva, dato che per avere informazioni significative per poter correlare un miglioramento in accuratezza ed alcuni indicatori di caratteristiche di immagini sarebbe necessario avere una popolazione di campioni di immagini diverse. Per quanto riguarda gli indicatori di caratteristiche, a parte le risoluzioni (spaziale, spettrale, radiometrica) si può tener conto anche della scala degli elementi da classificare (la dimensione media in pixel di una regione classificata), forme degli elementi, tessitura etc. Con l'utilizzo sempre più comune di procedure complesse per la classificazione, grazie alla disponibilità di algoritmi contextual, object-based ed anche quelli classici pixel based, sarebbe molto indicato fare una raccolta di dati di accuratezza dei risultati delle varie classificazioni, per poi avere un sistema informativo che permetta ad un operatore di scegliere il classificatore o i classificatori più adeguati.

## Conclusioni

E' evidente che determinati classificatori funzionino meglio in diverse circostanze. L'utilizzo dell'algoritmo Maximum Likelihood, ad esempio, rimane sempre la strategia migliore da adottare quando non si conoscono caratteristiche particolari delle classi oltre quelle riguardanti la riflettanza nelle zone dello spettro coperte dall'immagine. Se invece si hanno conoscenze più approfondite e quindi si possono utilizzare ulteriori elementi percettivi come fonte di differenziazione tra classi, allora si può procedere diversamente. La presenza di relazioni probabilistiche tra classi rende un procedimento come lo SMAP più indicato. Dato che la classificazione della copertura del suolo ha sempre una componente di relazioni spaziali probabilistiche tra classi, l'utilizzo dello SMAP porta quasi sempre ad un miglioramento, come è stato provato per le circostanze presenti nelle immagini in questa nota. L'unica situazione dove questo potrebbe non essere il caso potrebbe essere quando un'immagine presenta delle regioni molto "polverizzate" e di piccola estensione. In questo caso lo SMAP tende a "equalizzare", come potrebbe fare un filtro passa basso, la divisione in classi. Infatti esso tende ad accorpare piccole regioni in regioni più grandi. Questo solitamente è un vantaggio, ma non sempre.

Rimane da menzionare l'importanza di un utilizzo di indicatori di forma, a nostro parere molto utile per portare un valore aggiunto alla classificazione. La parte più complessa è esprimere matematicamente queste caratteristiche per poter utilizzarle come criterio decisionale. Un esempio è sicuramente l'utilizzo come indicatore di un coefficiente di forma che potrebbe essere espresso da un semplice rapporto tra perimetro ed area fino a più complessi indicatori. Questo potrebbe essere un criterio di assegnazione ulteriore che si va ad aggiungere al calcolo delle probabilità di assegnazione delle classi, dando al "fuzzy calculation" un elemento in più. Aumentando questi criteri, che, ovviamente andranno pesati a seconda della loro influenza nella classificazione, le procedure di classificazione andranno sempre più verso un vero e proprio sistema esperto. Questi sistemi incorporano diverse metodologie integrandole tra loro per fornire un supporto decisionale, nel nostro caso nella decisione di assegnazione ad una classe o ad un'altra. Lo svantaggio di utilizzare sistemi complessi è che essendo i nodi decisionali più numerosi, più numerose sono le potenziali fonti di errore, e quindi si richiede operatori con esperienza oltre la media per calibrare il sistema.

## **Bibliografia**

C. A. Bouman and M. Shapiro, "A Multiscale Random Field Model for Bayesian Image Segmentation," IEEE Trans. on Image Processing, vol. 3, no. 2, pp. 162-177, Marzo 1994.

R. de Kok, T. Schneider, U. Ammer, OBJECT-BASED CLASSIFICATION AND APPLICATIONS IN THE ALPINE FOREST ENVIRONMENT, International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 7-4-3 W6, Valladolid, Spain, 3-4 Giugno, 1999