

**ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA**

---

**FACOLTA' DI INGEGNERIA**

**CORSO DI LAUREA SPECIALISTICA IN INGEGNERIA CIVILE**

**D.I.S.T.A.R.T.**

Dipartimento di Ingegneria delle Strutture, dei Trasporti, delle Acque,  
del Rilevamento, del Territorio

**TESI DI LAUREA**

in

TELERILEVAMENTO LS

***ESTRAZIONE DI FEATURES CON TECNICHE DI  
CLASSIFICAZIONE  
PIXEL E OBJECT-ORIENTED***

CANDIDATO

Chiara Martelli

RELATORE:

Chiar.mo Prof. Ing. Gabriele Bitelli

CORRELATORI

Pof. Luca Vittuari

Ing. Alessandro Mognol

Anno Accademico 2008/09

Sessione III



# Indice

<b>Introduzione</b> .....	pag. 1
<b>1. Generalità</b> .....	” 5
<b>2. La classificazione</b> .....	” 11
2.1. La classificazione automatica.....	” 15
2.1.1 La classificazione supervisionata.....	” 17
2.1.2 La classificazione non supervisionata.....	” 29
2.2. Valutazione dell'accuratezza della classificazione.....	” 34
<b>3. Estrazione di Features</b> .....	” 37
3.1. La classificazione pixel-based.....	” 41
3.2. La classificazione object-based.....	” 42
3.2.1 La segmentazione.....	” 44
3.2.1 Segmentazione multirisoluzione.....	” 46
<b>4. Software utilizzati</b> .....	” 53
4.1. Envi 4.5.....	” 53
4.1.1 Envi Zoom: estrazione delle features.....	” 58
4.2. Definiens Developer 7.0.....	” 87
4.3. Considerazioni.....	” 96

5.	<b>Casi di studio</b> .....	”	97
5.1.	Conteggio di reperti archeologici di piccole dimensioni	”	98
5.1.1	Estrazione frammenti dall’immagine 1201.....	”	102
5.1.2	Estrazione frammenti dall’immagine 1219.....	”	109
5.1.3	Considerazioni sui risultati.....	”	114
5.2.	Estrazioni di edifici da immagini satellitari ad alta risoluzione.....	”	116
5.2.1	Primo caso di studio: Comune di Bologna.....	”	116
5.2.2	Secondo caso di studio: area di Caselle.....	”	119
5.2.3	Terzo caso di studio: Comune di Rimini.....	”	124
6.	<b>Conclusioni</b> .....	”	127
	<b>Bibliografia</b> .....	”	131





# **Introduzione**

Uno degli obiettivi primari nelle applicazioni del Telerilevamento satellitare è quello di estrarre dalle immagini informazioni tali da identificare le categorie di oggetti in esse presenti. L'obiettivo di questa tesi è quello di confrontare alcune tra le principali metodologie di analisi e classificazione di immagini tramite casi applicativi selezionati che possano evidenziarne pregi e difetti.

Le metodologie di classificazione si sono evolute nel tempo con lo scopo di superare limiti tecnici ed intrinseci agli algoritmi stessi, si sono sviluppati nuovi approcci e si è ampliato lo spettro delle applicazioni. Le immagini satellitari ad alta risoluzione spaziale si prestano per esempio ad essere utilizzate per analisi su larga scala di aree urbane, permettendo un livello di dettaglio elevato delle caratteristiche di questo ambiente. L'eterogeneità spaziale e spettrale di questa tipologia di aree e l'alta risoluzione del sensore rende molto complessa l'applicazione dei tradizionali algoritmi di analisi dell'immagini che in genere producono risultati non soddisfacenti. I motivi principalmente sono due: il primo è dovuto alle dimensioni modeste degli oggetti, il secondo è causato dall'eterogeneità spettrale interna a una stessa classe. La tecnica di classificazione pixel-based, la più frequentemente utilizzata, non risulta vincente sotto queste condizioni, e la più recente metodologia object-oriented potrebbe risultare più idonea.

Le tecniche di classificazione pixel-based fanno uso delle sole informazioni spettrali e permettono di utilizzare sia metodi non supervisionati che supervisionati. La metodologia non supervisionata non lascia molto spazio interpretativo all'utente: l'analista ha poco controllo sulle classi che vengono definite nell'immagine e il confronto tra istanti temporali diversi risulta difficoltoso perché il continuo cambiamento delle proprietà spettrali dei sensori, nel corso del tempo, causa variazioni nelle relazioni tra la risposta spettrale e la classe d'informazione che non rimangono perciò costanti. La tecnica di classificazione supervisionata in genere produce risultati che presentano una precisione superiore e permette agli utenti una possibilità d'intervento maggiore; ad esempio è lecito individuare e distinguere zone d'errore in cui le diverse regioni hanno lo stesso comportamento spettrale ma non lo stesso significato semantico (esempio: nevu-nevole). Nella classificazione supervisionata l'individuazione e il posizionamento delle varie tipologie di copertura del suolo sono noti a priori in alcune aree tramite rilievi sul campo, fotointerpretazione di immagini aeree, mappe, interpretazione visiva dell'analista: l'analista ricerca siti specifici che rappresentino esempi omogenei della copertura del terreno. Anche la classificazione supervisionata presenta dei limiti perché è fortemente dipendente dalla scelta dei training sites e dalla suddivisione dello spazio delle features. Si presentano problemi se si verificano sovrapposizioni tra classi o se alcuni pixels ricadono fuori dai clusters; questi fenomeni comportano un'errata classificazione dei dati e quindi è elevata la probabilità della presenza di errori di classificazione o pixel non classificati.

Studi recenti hanno dimostrato la superiorità dei metodi contestuali object-based rispetto ai concetti tradizionali di classificazione pixel-based. L'uso congiunto di informazioni quali forma, tessitura, informazioni contestuali, insieme a quelle spettrali permette una migliore caratterizzazione degli oggetti del mondo reale, avvicinandosi all'interpretazione visiva eseguita dall'occhio umano. La classificazione object-based inizia con un processo di segmentazione dell'immagine. Lo scopo principale della segmentazione è quello di identificare all'interno dell'immagine gli oggetti del mondo reale che descrivono il fenomeno

da studiare. La segmentazione di un'immagine è ottimale se gli image-objects che produce vanno a coincidere, o quasi, con gli oggetti del mondo reale. Dagli image-objects verranno calcolate le nuove features contestuali che verranno utilizzate nella fase successiva: la classificazione. È di facile intuizione che l'accuratezza della segmentazione è altamente correlata all'accuratezza della classificazione.

Con questa tesi si vuole in primo luogo verificare la flessibilità della segmentazione per l'estrazione delle classi e testare l'applicabilità del nuovo approccio di segmentazione multirisoluzione con differenti tipologie di dato. Nella prima parte della sperimentazione, legata ad una applicazione di interesse in campo archeologico, verranno analizzate immagini RGB non satellitari per estrarre elementi di piccole dimensioni, al fine di sviluppare una procedura in grado di effettuare un conteggio automatico delle classi d'interesse. Nella seconda parte sono invece prese in esame immagini ad alta risoluzione acquisite col satellite QuickBird che interessano aree urbane; in questo l'approccio object-based è rivolto all'estrazione di edifici.

La tesi è strutturata come segue.

Il primo capitolo è dedicato ad una trattazione generale sul Telerilevamento, dalla sua storia agli ambiti applicativi principali.

Nel secondo capitolo vengono descritte in generale le metodologie di classificazione delle immagini: le fasi che la contraddistinguono, la classificazione automatica, gli algoritmi disponibili, i criteri di validazione della classificazione.

Nel terzo capitolo sono trattati gli approcci pixel e object-based per l'estrazione di features; vengono evidenziati il metodo con cui operano, le differenze, i pregi, le informazioni che riescono a cogliere dalle immagini, la loro struttura interna. In particolare per il metodo object-based viene descritta la fase di segmentazione e, in dettaglio, l'algoritmo di segmentazione multirisoluzione.

Nel quarto capitolo vengono illustrati i programmi software adoperati nell'ambito della sperimentazione: si tratta di due tra i più importanti pacchetti professionali oggi disponibili.

Nel quinto capitolo viene descritta la sperimentazione e vengono riportati i risultati ottenuti.

Il sesto capitolo è dedicato infine ad alcune conclusioni sull'esperienza condotta.

# Capitolo I

## Generalità

Il Telerilevamento, chiamato in letteratura anglosassone Remote Sensing, è una disciplina tecnico-scientifica che permette di ricavare informazioni, qualitative e quantitative, sull'ambiente e su corpi posti a distanza, mediante misure di radiazione elettromagnetica che viene emessa, riflessa o trasmessa dalle superfici in oggetto. Queste misurazioni di radiazione elettromagnetica vengono fornite sottoforma di immagini, o per meglio dire dati numerici, e sono rilevate da sensori a bordo di aerei, satelliti o sonde spaziali.

Il telerilevamento consente studi con finalità diagnostico-investigative delle quali molteplici sono in ambito ambientale:

- Monitoraggio dell'atmosfera: composizione, profili meteorologici, precipitazioni, vento, input modelli di previsione e nowcasting (attività di previsione meteorologica a brevissimo termine pari alle tre ore successive), e dell'Oceano (comprese attività in mare e aree costiere): circolazione marina/oceanica, moto ondoso, livello del mare, temperatura, biologia marina, inquinamento acque;
- Climatologia: comprensione e valutazione dei processi climatici (ad esempio la desertificazione);
- Glaciologia: tipologia e movimenti dei ghiacci marini e terrestri;
- Geologia, geomorfologia, geodesia: faglie e anomalie, movimenti tettonici, instabilità pendii, identificazione tipi di roccia;

- Topografia e cartografia: carte geografiche di base e tematiche (es copertura del suolo);
- Agricoltura, foreste, botanica: mappatura, stato fenologico e di salute, umidità, stima di produzione, indicazioni colturali;
- Idrologia e idrogeologia: risorse idriche, precipitazioni, input modelli di run-off, evapotraspirazione, erodibilità dei suoli;
- Previsione e controllo catastrofi, valutazione dei rischi: incendi, alluvioni, instabilità, controllo zone inaccessibili;
- Gestione del territorio: inventari risorse. Supporto alla pianificazione, valutazione impatto ambientale.

La nascita e la messa a punto del Telerilevamento sono legate allo sviluppo tecnologico in queste tre discipline:

- fisica delle onde elettromagnetiche: oltre allo spettro visibile, l'infrarosso è stato scoperto nel 1800 da Herschel durante lo studio dell'effetto termico delle varie parti dello spettro elettromagnetico servendosi di termometri; l'ultravioletto è stato osservato, all'inizio del 1800 da Ritter grazie alla maggior reazione fotochimica che produce annerendo alcuni sali d'argento;

- tecniche fotografiche ed elettroniche: all'inizio del 1800 si sono manifestate una ricerca ed un affinamento tecnologico che hanno consentito la realizzazione di dispositivi capaci di rilevare e rappresentare le radiazioni; lo sviluppo dei sensori è stato incentivato durante il periodo bellico e sono stati effettuati studi sulle reazioni tra l'energia elettromagnetica ed i materiali fotosensibili;

- tecnologia aerospaziale: l'eccentrico fotografo Félix Tournachon detto il Nadar fu il primo ad appassionarsi alle innovazioni tecnologiche ed ebbe l'idea di attrezzare un pallone aerostatico con una camera fotografica nel 1856 per effettuare riprese stereoscopiche dall'alto. Lo stesso Nadar capì di aver dato il via ad un sistema di fotografia aerostatica che "permetteva di impiegare la fotografia per il rilievo di carte topografiche, idrografiche, catastali" utile anche per fini bellici

come “direzione di operazioni strategiche per individuare fortificazioni ed uomini in marcia”. Pochi anni dopo la nascita dell’aeromobile, esattamente nel 1909, i fratelli Wright eseguirono il primo rilievo aereo-fotografico per la produzione di cartografia. Le riprese aeree ebbero un grande impiego durante i conflitti mondiali come tecnica di rilevamento che di ricognizione, nel periodo post bellico, invece, la potenza sovietica e quella americana presero a gareggiare per la conquista dello spazio extra-atmosferico. Il primo aprile del 1960 ebbe inizio l’osservazione sistematica della Terra da distanze orbitali (da 700 a 800 km) tramite il satellite TIROS-1 munito di banda infrarosso per studi meteorologici; nel 1968 il programma Apollo fu il primo ad utilizzare il telerilevamento come strumento di osservazione e studio specifico delle risorse terrestri da distanze orbitali. Lo studio di Krinov sulle caratteristiche di riflettanza spettrale delle superfici naturali, pubblicato nel 1946, diede impulso allo sviluppo del Telerilevamento in campo civile; nel 1972 si ebbe il lancio del primo satellite della serie Landsat che convalidò questa tecnologia come uno strumento innovativo idoneo allo studio dell’habitat terrestre grazie alla qualità e al numero di immagini prodotte proponendo una nuova mentalità per l’approccio delle problematiche per lo studio dell’ambiente e del territorio.

Il Telerilevamento usa come veicolo d’informazione l’energia elettromagnetica in tutte le sue forme d’interazione con la superficie della Terra (emissione, riflessione, diffusione...) ed è costituito da cinque aspetti fondamentali, che sono:

- le caratteristiche della sorgente di energia che può essere di tipo naturale o artificiale; è un aspetto rilevante perché ne conseguono la fisica dell’interazione, la tipologia del sensore, l’analisi e l’interpretazione dei dati. Il Sole è la principale sorgente naturale di energia elettromagnetica;
- l’interazione con la superficie terrestre; ogni tipologia di superficie ha una differente risposta ma sussistono problemi di ambiguità nella sua interpretazione;

- l'interazione con l'atmosfera; il passaggio lungo l'atmosfera causa modifiche nella radiazione, sottoforma di diffusione e assorbimento che operano in modo diverso alle diverse lunghezze d'onda;

- tipologia di sensori e rilevatori: possono essere di tipo ottico-elettronico o sensori fotografici (emulsioni). Per sensore viene inteso qualunque dispositivo capace di rilevare e registrare l'energia elettromagnetica proveniente da una scena e convertirla in informazioni disponibili a successive analisi, ad esempio sottoforma di segnale elettrico. È necessario garantire misure confrontabili nel tempo, quindi con valori assoluti e riferimenti di calibrazione stabili a parità di costanza di funzionamento strumentale; ciò deve essere applicato nelle caratteristiche principali dei dati: geometriche, spettrali e radiometriche. I sensori vengono inoltre distinti in due categorie: quelli di tipo attivo in grado di produrre autonomamente l'energia elettromagnetica necessaria per riprendere la scena e quelli di tipo passivo ovvero progettati per rilevare l'energia elettromagnetica presente nell'ambiente originata da sorgenti naturali come le camere fotografiche e aerofotogrammetriche, fotocamere multispettrali;

- il processo di elaborazione ed interpretazione; in questo aspetto prevale la componente dell'operatore perché si avvale della conoscenza degli esperti multidisciplinari in grado di sviluppare macchine e sistemi capaci di elaborare una grande entità dei dati acquisiti in continuo, e sviluppare metodiche e modelli interpretativi. I computer vengono addestrati con programmi che tentano di simulare i processi mentali di *fotointerpretazione* per ottenere, a seguito di un processo del programma, risultati finali con caratteristiche di leggibilità diretta e perciò comprensibili alla percezione dell'uomo e compatibili con le sue esigenze.

La qualità dei dati di Telerilevamento è il risultato di un processo di trasferimento basato sulle specifiche tecniche degli apparati strumentali utilizzati durante le fasi di ripresa e di rappresentazione. Le specifiche di risoluzione sono tra le caratteristiche più importanti dei sistemi di osservazione, in quanto legate alla capacità di descrivere la scena rilevata in termini di dettaglio geometrico o spaziale, di contenuto spettrale e sensibilità radiometrica dei rilevatori, di

frequenza temporale o ripetitività del rilievo. La risoluzione geometrica esprime il dettaglio raggiungibile da un sistema di ripresa, viene espressa come la dimensione coperta al suolo da un singolo pixel ed è legata quindi alla capacità di discriminare in modo distinto due oggetti contigui. La risoluzione spettrale è legata alla capacità di rilevare su diverse porzioni dello spettro elettromagnetico per facilitare la riconoscibilità fisica della superficie oggetto di studio. La risoluzione radiometrica è legata al processo di discretizzazione del dato radiometrico e in ultima analisi al numero di livelli di grigio impiegati nella rappresentazione della scena: se è elevata la rappresentazione della scena è più fedele e quindi anche la capacità di estrarre da essa le informazioni. La risoluzione temporale è legata alla possibilità di ripetere le osservazioni a cadenze temporali costanti o programmabili. È un aspetto basilare nelle casistiche di controllo della dinamica dei fenomeni naturali perché vincola la disponibilità dei passaggi utili, anche se il numero di immagini finali risulta di norma inferiore a causa delle condizioni dell'atmosfera (es. presenza di nubi durante l'acquisizione).

Ad oggi non è ancora possibile avere contemporaneamente le massime prestazioni per ogni tipologia di risoluzione; scegliere un'alta risoluzione geometrica o spettrale costringe ad una bassa risoluzione temporale, invece richiedere un'alta risoluzione radiometrica implica una scarsa risoluzione geometrica e spettrale; in genere è buona norma cercare un compromesso tra risoluzione spettrale e geometrica per garantire un buon rendimento radiometrico.

Il telerilevamento offre un'informazione globale sull'ambiente ed il territorio presentando interessanti caratteristiche quali:

- punto di vista dell'osservatore non convenzionale, che consente una relativa facilità ad ottenere informazioni di qualsiasi tipo su vaste aree, in particolare per territori di difficile accesso;
- carattere sinottico e coerenza delle informazioni;
- rapida acquisizione delle informazioni (anche in tempo reale) ed in minor tempo rispetto alle tecniche tradizionali;

- ripetitività nel tempo del rilievo;
- investigazione in svariate bande dello spettro elettromagnetico senza perturbare l'oggetto del rilievo.

Nonostante questi aspetti positivi vi sono alcuni inconvenienti tipici che possono peggiorare i dati come le ombre e le nubi che rendono più complessa l'analisi dell'immagine: per questo è necessaria una pre-elaborazione dell'immagine, con la quale vengono applicate correzioni geometriche e radiometriche, ed inoltre è sempre indispensabile una campagna di validazione dei dati a terra per una corretta georeferenziazione.

Molte di queste considerazioni si applicano alle immagini nel campo ottico ma non a quelle nel campo radar.

## Capitolo II

# La classificazione

I dati da telerilevamento derivano dal trasferimento delle informazioni dal mondo reale ad una realtà oggettiva che viene rappresentata nelle immagini. Le immagini descrivono un fenomeno, oppure un processo, tramite una rappresentazione, in genere bidimensionale, delle informazioni relative al fenomeno o al processo di studio. La rappresentazione della scena osservata è legata ad un insieme di informazioni relative alle interazioni della radiazione elettromagnetica con le superfici fisiche. Nota la tipologia del dato e conoscendo a priori il processo di formazione, le caratteristiche e le tecniche di elaborazione che lo contraddistinguono, diventa necessaria una chiave interpretativa per l'estrazione in modo qualitativo e quantitativo delle informazioni.

I processi di interpretazione sono sostanzialmente distinguibili in due tipologie: il processo d'interpretazione qualitativo tipico della mente umana e il processo di interpretazione quantitativo tipico delle macchine. Come può essere di facile intuizione questi approcci risultano completamente differenti. Il primo, che fa riferimento sostanzialmente alla disciplina della fotointerpretazione, è interamente realizzato da un esperto analista che attua un processo decisionale ad alto livello servendosi con facilità delle informazioni spaziali a disposizione (forma, dimensione e tessitura), ma risulta limitato nell'analisi multispettrale e nella distinzione dei livelli di grigio e dei toni di colore, produce in genere una scarsa accuratezza nel calcolo di dati metrici ed è un metodo soggettivo, concreto, qualitativo. Il secondo, legato all'interpretazione automatica, è caratterizzato da un processo decisionale a basso livello, una limitata iterazione con l'operatore umano,

non eccelle nella comprensione delle informazioni spettrali ma consente una buona analisi multispettrale, un'analisi quantitativa di tutti i livelli di grigio o dei toni di colore, una valutazione precisa delle dimensioni, ed è un metodo oggettivo, astratto e quantitativo.

L'interpretazione visiva delle immagini è in funzione di un complesso processo di tipo analogico: per la mente umana è legata alla capacità di adottare delle generalizzazioni, ovvero passare da singole informazioni ad un modello astratto: risultano per questo necessarie conoscenze pregresse che consentano l'identificazione di oggetti facendo uso delle loro caratteristiche di forma, dimensione, colore, tessitura, struttura, ombra e ubicazione. Il trattamento dei dati telerilevati, anche in considerazione della loro entità e della complessità nella loro gestione, ha reso obbligatorio lo sviluppo di procedure automatiche di interpretazione per contenere i tempi di elaborazione entro limiti ragionevoli ma comunque capaci di fornire informazioni complete sul caso di studio. L'uso di procedure automatiche permette di standardizzare i criteri adottati e di rendere possibili controlli quantitativi sull'accuratezza dei risultati.

L'analisi di un'immagine è finalizzata ad estrarre il contenuto informativo in modo comprensibile all'utente finale. Una delle informazioni richieste può essere un'identificazione delle categorie o classi presenti nell'immagine oggetto di studio. La *pattern recognition*, una sottoarea dell'apprendimento automatico, cerca di emulare il processo tipico della mente umana; tale disciplina studia sistemi per il riconoscimento in automatico di oggetti (identificazione di *pattern* ovvero modelli teorici) all'interno dei dati al fine della classificazione. Durante il processo d'identificazione si ha un'analisi che ha l'obiettivo di individuare gli elementi appartenenti ad una classe in modo tale da rispettare criteri di similarità o associazione elevati tra gli elementi appartenenti alla stessa categoria.

Con la procedura di classificazione delle immagini telerilevate vengono associate al pixel informazioni di tipo semantico relative ad uno specifico aspetto o classe. La classificazione automatica è l'alternativa moderna alla fotointerpretazione; è caratterizzata da un processo astratto univocamente definito

in modo formale, quindi ripetibile e quantificabile. Le differenze sostanziali risiedono nei diversi modi di operare: l'analista lavora su porzioni dell'immagine considerando le interrelazioni tra i pixels, la classificazione automatica opera nello spazio multi-dimensionale delle caratteristiche.

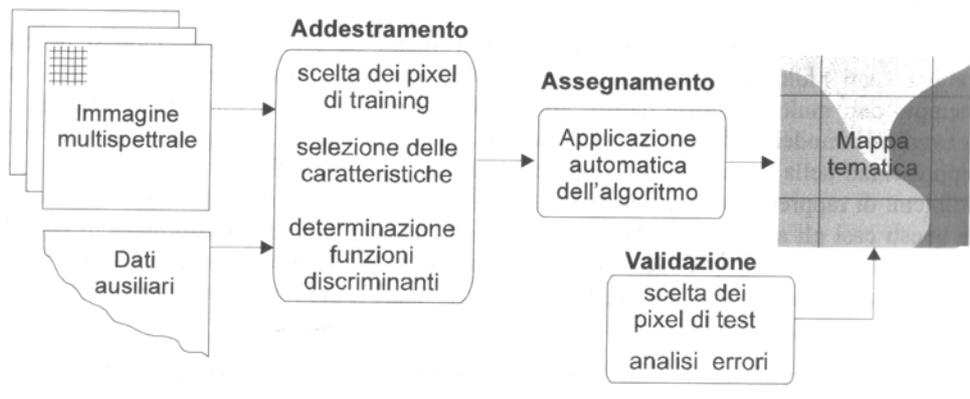
Lo spazio delle caratteristiche è uno spazio  $n$  dimensionale, dove  $n$  è pari al numero di informazioni a disposizione per ogni pixel; esso coincide con il numero di bande spettrali. Le features<sup>1</sup> sono disponibili quando lo studio non è al livello del singolo pixel ma viene eseguito su objects ricavati da una segmentazione. In caso di features non omogenee è opportuna una normalizzazione dimensionale, ad esempio se devono essere confrontati un perimetro ed un'area, nello spazio delle features, che è un piano, al posto dell'area verrà considerata la sua radice quadrata per omogeneizzare le unità di misura.

La classificazione è un processo suddiviso in due passaggi: il primo dallo spazio geografico relativo all'immagine acquisita allo spazio delle caratteristiche relativo alle informazioni misurate, il secondo dallo spazio delle caratteristiche allo spazio geografico dell'immagine classificata (mappa tematica) corredata dalle informazioni semantiche delle classi.

---

<sup>1</sup> Una nota sulla terminologia, emersa dall'esperienza di tesi: occorre osservare che nella Geomatica il termine "feature" viene associato a concetti diversi a seconda della disciplina, ed anche all'interno di uno stesso ambito possono trovarsi connotazioni diverse. E' per esempio differente il significato che si attribuisce al termine "features" nella fotogrammetria digitale e nelle applicazioni orientate alla robotica (es. nella correlazione automatica di immagini) da quello che si intende nell'analisi radiometrica di un'immagine multispettrale (con spazio delle feature ci si riferisce sostanzialmente ad uno scatterogramma a  $k$  dimensioni, dove  $k$  è il numero di bande) o infine da quello che si ritrova nei pacchetti software per la classificazione automatica di immagini. In quest'ultimo caso si ritrova lo stesso termine associato sia agli oggetti estratti (es. software ENVI Feature Extraction) che agli attributi (spaziali, tessiturali, ecc.) degli oggetti stessi (software eCognition).

Schema delle principali fasi del processo di classificazione.



La differenziazione degli elementi in classi distinte avviene mediante metodi matematici o statistici che consentono un'analisi di alcune proprietà o caratteristiche, espresse in forma numerica, degli oggetti nello spazio delle caratteristiche. L'assegnazione di una classe dipende da regole di decisione che sono composte da funzioni di decisione, chiamate anche funzioni discriminanti, a valori reali definite sul vettore delle misure che devono essere definite per ogni classe; nel contesto della teoria statistica, corrispondono alle funzioni densità di probabilità, nel contesto degli insiemi fuzzy sono funzioni di appartenenza. Chiamando  $d_i(\underline{x}) = d_i(x_1, x_2, \dots, x_k)$  la funzione discriminante caratterizzata dal vettore di misure  $\underline{x} = (x_1, x_2, \dots, x_k)$ , la classe  $\omega_i$  viene assegnata all'oggetto per il quale la funzione discriminante è massima:

$$\underline{x} \in \omega_i \Rightarrow d_i(\underline{x}) > d_j(\underline{x}) \forall j \neq i$$

Questa operazione suddivide lo spazio delle caratteristiche  $k$ -dimensionale in regioni di decisione, che coincidono alle classi (classi informative), delimitate da una frontiera chiamata decision boundary; in questo modo è possibile decidere quale classe assegnare al campione sulla base della posizione del punto nel feature space. La separazione tra classi è effettuata sulla base delle conoscenze delle proprietà fisiche, delle risposte spettrali delle superfici e sul concetto di similarità tra pixels. I pixels appartenenti alla stessa tipologia di oggetto sul terreno, quindi

della stessa classe informativa, tendono ad raggrupparsi attorno certi punti del feature space formando nuvole (cluster) che rispecchiano le classi spettrali; la condizione ideale sarebbe quella in cui un cluster rispecchia una classe informativa, quindi sia presente un legame univoco tra classe spettrale e classe d'informazione. Tanti fattori come la variabilità atmosferica, la situazione topografica, diverse condizioni d'illuminazione e ombra, firme spettrali, pixels misti, comportano variazioni spettrali delle classi d'informazione; può accadere che la stessa classe d'informazione sia costituita da più classi spettrali che presentano un certo grado di separazione tra loro, quindi si presenta una variabilità all'interno della stessa classe che soltanto in certi casi può essere modellata.

La classificazione di un'immagine multispettrale, risulta tanto più accurata quanto il grado di separazione delle classi, nello spazio delle caratteristiche, è elevato; inoltre il livello di dettaglio riconoscibile nell'immagine e la scelta delle classi tematiche risultano condizionanti in questa fase.

### **2.1 La classificazione automatica**

Le caratteristiche del sensore e della ripresa influenzano la suddivisione in pixels del territorio. La realtà continua viene frammentata in unità spaziali, si verificheranno due condizioni: una in cui gli elementi saranno omogenei (pixels puri) e quindi ciò che rappresentano appartiene alla stessa categoria, un'altra in cui gli elementi non saranno omogenei (pixels misti) perché al loro interno sono presenti diverse classi. Convenzionalmente le tecniche di classificazione operano supponendo che ogni pixels dell'immagine rappresenti un'unica categoria, l'assegnazione della classe avviene con logica booleana di vero o falso; questo tipo di classificazione viene detta hard ed è contrapposta ai metodi non convenzionali (classificazione soft) che diversamente operano associando ad ogni pixels più di una classe corredata dal grado di apparenza della classe secondo i concetti di logica multivalore come nel caso degli insiemi fuzzy.

Il processo di classificazione è costituito da tre passi fondamentali che devono essere eseguiti secondo questo ordine: addestramento, assegnamento e validazione.

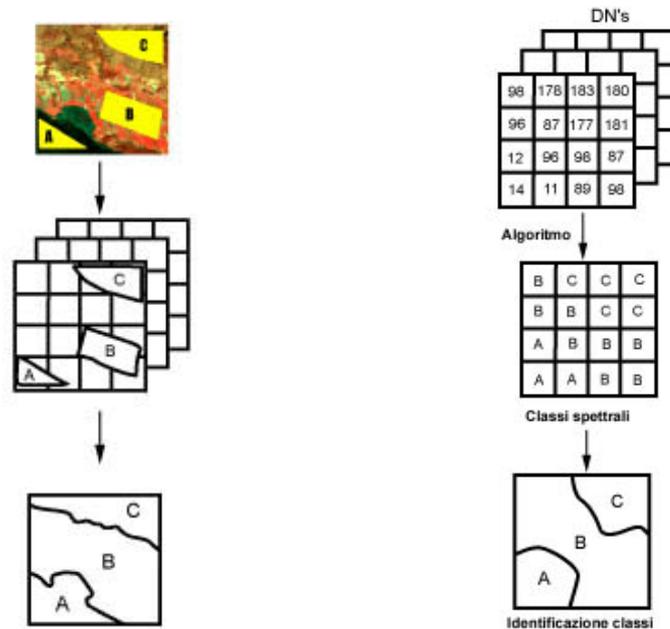
Durante la fase di addestramento, detta anche di training, l'esperto deve identificare un'insieme di pixels campioni (tra 1% e 5% dei pixels totali) idonei a rappresentare e caratterizzare ciascuna delle classi d'interesse. Nella pratica questi pixels vengono individuati tramite regole di decisione, applicate allo spazio delle caratteristiche, che vengono estese a tutti i pixels che diventano così classificati nella fase di assegnamento. La validazione serve per dare una valutazione del risultato della classificazione automatica; tramite un'analisi degli errori viene verificata la capacità di generalizzazione del classificatore. L'accertamento dell'accuratezza finale dell'immagine prodotta coinvolge l'identificazione un set di ubicazioni campione che sono presenti nella realtà, che vengono poi comparate poi con ciò che si prevedeva nell'immagine per la stessa ubicazione.

Ci sono due approcci generali per la classificazione dell'immagine: supervisionata e non supervisionata.

Nel caso della classificazione supervisionata, il compito è scoprire specifici tipi di copertura di terra, definiti a priori dall'utente tramite l'individuazione di dati di training, ovvero pixels rappresentativi per ogni tipologia di classe. In questo modo il programma viene addestrato a riconoscere le peculiarità per ogni classe.

Nella classificazione non supervisionata non sono definite a priori le classi di interesse quindi non sono richieste informazioni di anticipo sulle categorie. Il programma esamina i dati ed li suddivide nei naturali raggruppamenti spettrali più comuni, cluster, presenti nei dati. L'analista identifica a posteriori questi cluster come classi di copertura del suolo, è importante ricordare che i cluster prodotti da questa tipologia di classificazione non sono classi d'informazioni, ma classi spettrali (raggruppano insieme pixels con modelli di riflettanza simili).

Sotto: a destra procedimento della classificazione supervisionata, a sinistra procedimento della classificazione non supervisionata



### 2.1.1 Classificazione supervisionata

La classificazione supervisionata è la tecnica più utilizzata per un'analisi quantitativa delle immagini da Telerilevamento. Può essere schematizzata nei seguenti step:

1. definizione della legenda: fase in cui deve essere specificato l'insieme delle classi di copertura presenti sul terreno da individuare sull'immagine;
2. scelta dei training samples: per ogni classe da identificare devono essere individuati dei pixels per l'addestramento del programma che eseguirà la classificazione;
3. selezione delle caratteristiche: vanno identificate le componenti del feature space tra tutte quelle disponibili all'interno dei dati di training;

4. funzioni di decisione: ad ogni classe viene associata regola di decisione per suddividere il feature space in regioni corrispondenti alle classi da assegnare;
5. applicazione dell'algoritmo: tramite un processo automatico tutti i pixels dell'immagine vengono classificati secondo le regole di decisione, il risultato finale è una mappa tematica in cui compaiono le classi richieste;
6. validazione dei risultati: si valuta l'accuratezza della mappa tramite un confronto del risultato con un insieme di pixels noti detti ground truth; il grado di confidenza dei risultati è rappresentato nella matrice di confusione.

### **La fase d'addestramento**

La fase di addestramento è quella che richiede il maggior tempo d'intervento dell'esperto che deve far ricorso alle proprie capacità, conoscenza del problema e del territorio per individuare i training samples più idonei a seconda delle circostanze e degli obiettivi dello studio; ad esempio una oggetto nel territorio come una coltivazione a grano può rientrare nella classe "vegetazione" oppure "coltura agricola", oppure "grano" fino ad arrivare alla distinzione tra "grano tenero" e "grano duro" in caso di studi specifici.

I training samples vengono ricercati con sopralluoghi in campo se il terreno è facilmente accessibile oppure, se non lo è, tramite immagini a risoluzione maggiore come fotografie aeree a grande scala o con l'aiuto di tecniche d'interpretazione visuale.

La scelta dei training samples deve essere accurata perché, tramite le informazioni che contengono, verranno estratte le firme spettrali delle varie classi; è buona norma operare selezionandoli in aree omogenee, rimuovendo gli elementi outlier che escono dal cluster della classe nello spazio delle caratteristiche, evitando aree di passaggio tra classi che presentano pixels misti, selezionandoli in

zone diverse dalle stessa immagine, per considerare la variabilità propria della classe, e in numero adeguato (in caso di tecniche di classificazione che fanno uso di parametri statistici sono necessari almeno  $k+1$  elementi in uno spazio delle caratteristiche  $k$ -dimensionale affinché la matrice di covarianza non sia singolare).

### **Selezione delle caratteristiche e separabilità tra classi**

Lo spazio delle caratteristiche comprende tutte le bande a disposizione dai dati ma è conveniente ricercare al suo interno un sottoinsieme di bande, anche sintetiche, che meglio differenzia le classi in modo da incrementare l'accuratezza della classificazione. Questa osservazione è comprovata dal fenomeno di Hughes che enuncia: l'accuratezza di un modello cresce inizialmente all'aumentare delle variabili (es. bande) finché non verrà raggiunto un limite massimo d'accuratezza oltre il quale l'accuratezza decresce al crescere delle variabili.

I metodi possibili per ridurre le dimensioni dello spazio delle caratteristiche, che consentono di ottenere il migliore sottoinsieme di informazioni, sono l'analisi delle componenti principali e l'analisi di separabilità.

L'analisi di separabilità valuta l'errore atteso nella classificazione considerando le diverse possibili selezioni di caratteristiche, viene condotta sui training samples e non sull'intera immagine. La similarità tra classi viene valutata valutando la distanza che presentano nello spazio delle caratteristiche, considerando due classi  $a$  e  $b$  in uno spazio delle caratteristiche  $k$ -dimensionale, le misure di distanza che possono essere utilizzate sono:

- $L_1$  o *city block*:  $L_1 = \|\mu_a - \mu_b\| = \sum_{k=1}^K |m_{ak} - m_{bk}|$
- $L_2$  o *Euclidea*:  $L_2 = \|\mu_a - \mu_b\| = \left[ (\mu_a - \mu_b)^T (\mu_a - \mu_b) \right]^{1/2} = \sqrt{\sum_{k=1}^K (m_{ak} - m_{bk})^2}$
- *Angolare*:  $ANG = \ar \cos \left( \frac{\mu_a^T \mu_b}{\|\mu_a\| \|\mu_b\|} \right)$

Queste distanze non considerano le sovrapposizioni delle distribuzioni delle classi perché sono basate unicamente sul valore medio della classi; per considerare la varianza  $\sigma$  e la covarianza  $C$  si possono utilizzare le distanze:

- $L_1$  normalizzata:  $NL_1 = \sum_{k=1}^K \frac{|m_{ak} - m_{bk}|}{\frac{1}{2}(\sqrt{\sigma_{ak}} + \sqrt{\sigma_{bk}})}$
- Mahalanobis:  $MH = \left[ (\mu_a - \mu_b)^T \left( \frac{C_a + C_b}{2} \right)^{-1} (\mu_a - \mu_b) \right]^{1/2}$

Queste distanze risultano nulle quando le due classi presentano lo stesso valore medio, perciò vengono preferite distanze che sono basate su distribuzioni di probabilità normali:

- Divergenza:
 
$$D = \frac{1}{2} \text{tr}[(C_a - C_b) \cdot (C_a^{-1} - C_b^{-1})] + \frac{1}{2} \text{tr}[(C_a^{-1} - C_b^{-1}) \cdot (\mu_a - \mu_b) \cdot (\mu_a - \mu_b)^T]$$
- Bhattacharyya:  $B = \frac{1}{8} MH + \frac{1}{2} \ln \left[ \frac{\left| \frac{C_a + C_b}{2} \right|}{\sqrt{|C_a| \cdot |C_b|}} \right]^{1/2}$

dove  $tr$  rappresenta la traccia delle matrici.

Anche questa tipologia di distanze presenta degli inconvenienti perché non sono limitate superiormente e crescono indefinitamente al crescere della separazioni tra le classi. Questo problema non è presente in queste distanze:

- Divergenza trasformata:  $DT = 2(1 - e^{-D/8})$
- Jeffries Matusita:  $JM = [2(1 - e^{-B})]^{1/2}$

L'introduzione dell'esponenziale fa sì che per questa tipologia di distanza il termine tra parentesi decresca al crescere della separazione tra le medie delle classi. Il valore del limite superiore delle distanza di Jeffries Matusita è pari a 2 e, nel caso di due sole classi, l'accuratezza risulta del 100%.

## Funzioni di decisione con approccio Bayesiano

Thomas Bayes ha messo a punto un approccio matematico idoneo a risolvere i problemi d'inferenza in cui è richiesto operare delle scelte in presenza di informazioni incomplete o incertezza, proprio come nel caso della classificazione. L'inferenza bayesiana fornisce strumenti logici per discriminare tra ipotesi alternative/in conflitto. Tramite tale approccio, si usa una stima del grado di confidenza in una data ipotesi prima dell'osservazione dei dati, al fine di associare un valore numerico al grado di confidenza in quella stessa ipotesi successivamente all'osservazione dei dati. Questo approccio è basato sull'assunto che le quantità in gioco sono governate dalle distribuzioni di probabilità e quindi le decisioni e le scelte ottimali vengono prese valutando la probabilità e i dati osservati; questo metodo è in grado di fornire strumenti quantitativi che mettono in luce le realtà che sopportano le differenti ipotesi tra cui operare la scelta. L'inferenza bayesiana è un approccio all'inferenza statistica in cui le probabilità non sono interpretate come frequenze, ma come livelli di fiducia nel verificarsi di un dato evento. Il conflitto insito tra classi non viene eliminato con completa certezza e gli errori saranno superiori quanto più elevata è la sovrapposizione delle funzioni di distribuzione.

Per poter operare con questo approccio è necessaria la conoscenza a priori di molte distribuzioni di probabilità, se non tutte sono note vengono stimate sulla base di conoscenze e dati (training samples) preesistenti, facendo ipotesi sulla curva di distribuzione di probabilità che in genere viene considerata normale.

## Algoritmi di decisione

Vengono illustrati di seguito gli algoritmi più utilizzati durante un processo di classificazione supervisionata di immagini multispettrali.

### 1. Algoritmo minima distanza dalla media (MDM)

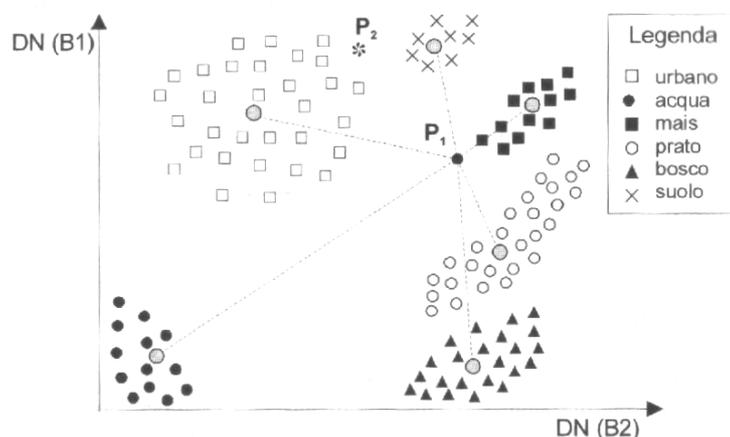
questo algoritmo fa uso soltanto del valore medio dei pixels di training per ognuna delle classi. Viene calcolata la distanza del vettore di misura  $x$  di ogni pixels dal

valore medio  $\mu_i$  di ogni classe  $\omega_i$ ; al pixel verrà assegnata la classe  $\omega_i$  per la quale la distanza  $d(\underline{x})$  risulta minima:

$$\text{il pixel } \underline{x} \in \omega_i \text{ se } d_i(\underline{x}) < d_j(\underline{x}) \text{ per ogni } j \neq i$$

i limiti decisionali della minima distanza rispetto la media sono in relazione alla funzione distanza utilizzata (distanza Euclidea, City Block...).

Questo metodo è matematicamente semplice e computazionalmente efficiente ma presenta il grosso limite di non riconoscere il diverso grado di dispersione delle classi nello spazio delle caratteristiche e non è nemmeno idoneo in quei casi in cui le classi sono molto ravvicinate.



Algoritmo MDM in uno spazio delle caratteristiche bidimensionale

## 2. Algoritmo del parallelepipedo

In questo caso le classi vengono definite da dei limiti inferiori e superiori. Questo algoritmo, per ogni classe, è in grado di stimare la dispersione nello spazio delle caratteristiche tramite degli intervalli, definiti sui pixels di training che sono stabiliti dal valore massimo e minimo delle bande spettrali. Un pixel verrà quindi assegnato se il suo valore è compreso in un certo range:

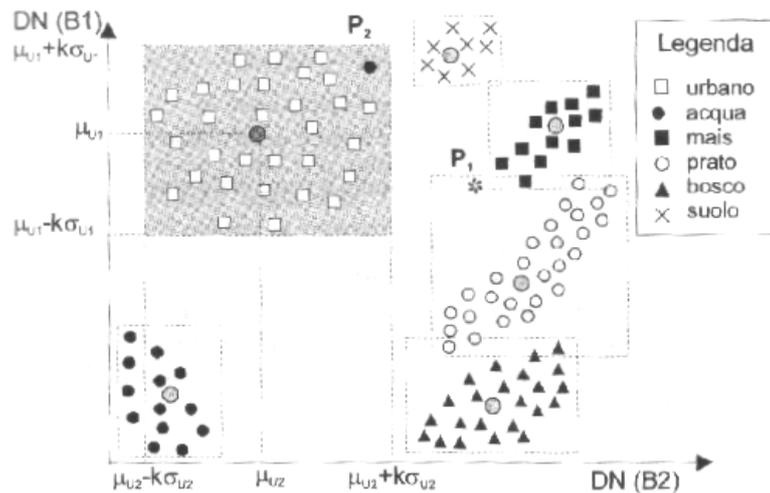
$$\text{il pixel } \underline{x} \in \omega_i \text{ se } Inf_i \leq \underline{x} \leq Sup_i \text{ per } i = 1, \dots, K$$

Se il pixel non soddisfa il criterio sarà assegnato alla classe *non classificato*.

Per ogni banda si avrà un intervallo, quindi nello spazio delle caratteristiche si formeranno dei parallelepipedi a  $n$  dimensioni. Un altro criterio possibile è

individuare i parallelepipedi facendo uso dei parametri statistici, media e deviazione standard, dei pixels di training:

il pixel  $x \in \omega_i$  se  $Inf_i \leq x \leq Sup_i$  dove  $Inf_i$  e  $Sup_i = \mu_i \pm \sigma_i$



Algoritmo del parallelepipedo in uno spazio delle caratteristiche bidimensionale

### 3. Massima verosimiglianza con modello gaussiano

Si basa sull'ipotesi che ogni classe  $\omega_i$  sia caratterizzata da una funzione distribuzione di probabilità di tipo gaussiano; se questa ipotesi è lecita l'algoritmo può essere utilizzato e la classificazione avviene secondo la regola di decisione bayesiana per la quale i pixels vengono assegnati alla classe che presenta la probabilità a posteriori maggiore; in questo modo si ha anche una suddivisione in classi che minimizza la probabilità media di errore sui dati. Per procedere sarà necessario settare i parametri: media, varianza e covarianza, della distribuzione di ogni classe, quindi è necessaria una conoscenza delle caratteristiche delle classi a priori.

Per ricavare la regola di assegnazione delle classi si parte dalla probabilità condizionale  $P(x | \omega_i)$ , o distribuzione densità di probabilità, della classe  $\omega_i$ , nello spazio delle caratteristiche uni-dimensionale o uni-variato, che è espressa in

$$\text{questo modo: } P(x | \omega_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$$

Invece che massimizzare la probabilità condizionale, si preferisce massimizzare il suo logaritmo che sicuramente sarà sempre una funzione monotona. La regola di decisione utilizzata nell'implementazione dell'algoritmo è la seguente:

$$D_i(x) = \ln[P(x | \omega_i)P(\omega_i)] = \ln[P(\omega_i)] - \frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_i^2) - \frac{(x - \mu_i)^2}{2\sigma_i^2}$$

Si fa notare che soltanto l'ultimo termine deve essere calcolato per ogni pixel perché in funzione di  $x$ .

È di interesse maggiore esprimere la regola di decisione per uno spazio  $k$ -dimensionale, caso delle immagini multispettrali, in cui la probabilità condizionata  $P(x | \omega_i)$  diventa una funzione multi variata  $k$ -dimensionale ed è espressa in questo modo:

$$P(x | \omega_i) = \frac{1}{(2\pi^{k/2}) \cdot |C_i|^{1/2}} e^{-\frac{1}{2}(x - \mu_i)^T C_i^{-1} (x - \mu_i)}$$

Dove :  $\mu_i$  è il vettore dei valori medi delle misure nella classe  $\omega_i$  ;

$C_i$  è la matrice di varianza e covarianza ( $k \times k$ ) della classe  $\omega_i$  ;

$|C_i|$  è il determinante della matrice  $C_i$  ;

$C_i^{-1}$  è la matrice inversa della matrice  $C_i$

I limiti di decisione saranno delle ipersuperfici, il luogo dei punti che presenta densità costante è dato da delle iper-ellissoidi che hanno la seguente espressione in forma quadratica:

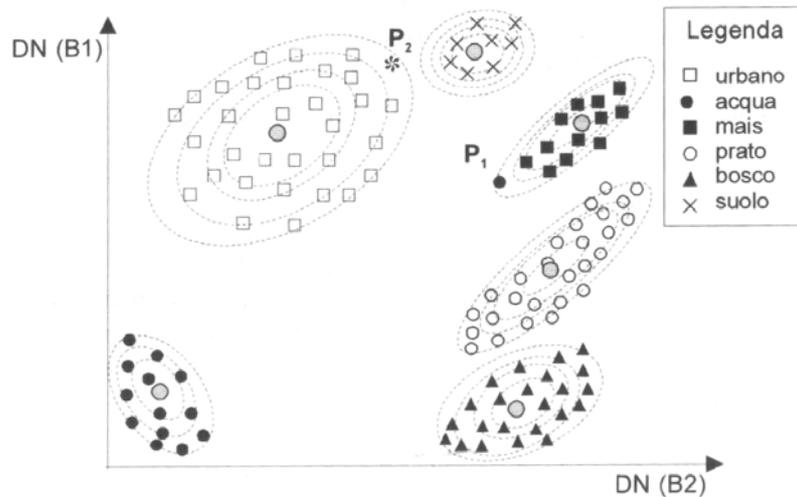
$$(x - \mu_i)^T C_i^{-1} (x - \mu_i) = r^2$$

che si mantiene costante. Il termine  $r$  è detto *distanza di Mahalanobis* tra  $x$  e  $\mu_i$ .

La regola di decisione viene così espressa:

$$D_i(x) = \ln[P(x | \omega_i)P(\omega_i)] = \ln[P(\omega_i)] - \frac{K}{2} \ln(2\pi) - \frac{1}{2} \ln(|C_i|) - \frac{1}{2} (x - \mu_i)^T C_i^{-1} (x - \mu_i)$$

Si nota come il criterio di decisione per assegnare un pixel ad una classe, non consideri soltanto la sua posizione  $x$  rispetto i valori medi delle classi, ma anche la dispersione della classe nello spazio delle caratteristiche. Gli ellissoidi di equiprobabilità ottenuti con questo metodo esprimono la dispersione delle classi.



Algoritmo della massima verosimiglianza in uno spazio delle caratteristiche bidimensionale

L'algoritmo calcola la verosimiglianza di appartenenza ad ognuna delle classi  $\omega_i$ , per ogni singolo pixel  $x$ , che verrà assegnato alla classe che presenta la verosimiglianza massima. Deve essere fissato un valore di soglia  $S_i$ , significativo per la classe spettrale  $\omega_i$ ; se i tutti valori di probabilità di appartenenza a posteriori risultano al di sotto della soglia  $S_i$ , i pixels verranno assegnati alla classe "non classificati".

Il valore di soglia viene determinato generalmente in questo modo:

$$\ln[P(\omega_i)] - \frac{K}{2} \ln(2\pi) - \frac{1}{2} \ln(|C_i|) - \frac{1}{2} (x - \mu_i)^T C_i^{-1} (x - \mu_i) > S_i$$

che può essere riscritta:

$$(x - \mu_i)^T C_i^{-1} (x - \mu_i) < 2S_i - \ln(|C_i|) + 2 \ln[P(\omega_i)]$$

Avendo assunto che  $x$  segua una distribuzione normale, il termine  $(x - \mu_i)^T C_i^{-1} (x - \mu_i)$  è contraddistinto da una distribuzione di tipo  $\chi^2$  a  $k$  gradi di libertà, pari alla dimensione dello spazio delle caratteristiche. Il livelli di confidenza, espressi in percentuale, ed i valori critici della distribuzione  $\chi^2$  possono essere determinati e vengono riportati e titolo d'esempio nella tabella sottostante al variare dei gradi di libertà:

Livello di confidenza	10%	5 %	2.5%	1%	0.5%
Gradi di libertà 3	6.25	7.81	9.35	11.34	12.84
Gradi di libertà 4	7.78	9.49	11.14	13.28	14.86
Gradi di libertà 6	10.64	12.59	14.45	16.81	18.55

#### **4. Distanza angolare spettrale – SAM (Spectral Angle Mapper)**

Questo algoritmo consente un metodo rapido di classificazione e viene spesso utilizzato per immagini iperspettrali che sono caratterizzate da un elevato numero di bande. Viene effettuata un'analisi di similarità tra gli spettri di riferimento delle classi e gli spettri associati ai pixels dell'immagine; questi spettri vengono trattati come dei vettori nello spazio delle caratteristiche. Gli spettri di riferimento sono ottenuti dalla media dagli spettri forniti dai training samples di ogni classe  $\omega_i$ . Il concetto di similarità è espresso come distanza angolare tra i due vettori degli spettri nello spazio n-dimensionale. La classe che viene assegnata è quella che presenta la distanza angolare minore, il pixel rimarrà non classificato se tale distanza supera una soglia fissata a priori. La distanza angolare  $\alpha$  è misurata in radianti, assume valori compresi tra zero e  $\pi/2$ , piccoli valori di  $\alpha$  indicano un'alta similarità nei confronti dello spettro di riferimento relativo ad una classe, alti valori invece rappresentano forti dissimilarità. L'angolo  $\alpha$  viene calcolato con un'espressione ricavata dalla definizione di prodotto scalare di due vettori. Dati due vettori  $u$  e  $v$  il loro prodotto scalare può essere così espresso:

$$u \cdot v = \sum_{i=1}^n (u_i v_i) = uv \cdot \cos \alpha$$

Dove  $u_i, v_i$  sono le componenti dei due vettori nello spazio n-dimensionale e  $u, v$  indicano i moduli dei due vettori. Mettendo in evidenza l'angolo  $u_i v_i$  si ha:

$$\alpha = \arccos \frac{u \cdot v}{uv} = \arccos \left( \frac{\sum_{i=1}^n (u_i v_i)}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}} \right)$$

## 2.1.2 Classificazione non supervisionata

Questa tipologia di classificazione opera sfruttando il concetto che oggetti simili producono risposte spettrali simili, non sono richieste conoscenze a priori sull'entità, tipologia, descrizioni delle classi, che vengono dedotte facendo osservazioni sui clusters che si formano nello spazio delle caratteristiche.

La peculiarità principale è che le classi vengono identificate da cluster compatti e ben distinguibili tra loro senza conoscere a priori l'entità della loro natura, l'intero spazio delle caratteristiche viene suddiviso in classi spettrali adoperando criteri di prossimità o similarità. Soltanto una volta terminato il processo d'identificazione delle classi spettrali l'analista assocerà a queste le classi tematiche in relazione alla conoscenza del territorio.

### Tecniche di clustering

Con le tecniche di clustering, i vettori misura associati ai pixels vengono assegnati ai clusters tramite regole di decisione basate sul concetto di similarità.

Come criterio di similarità vengono in genere preferite misure di distanza; indicando con  $\underline{x}$  il vettore delle misure spettrali associate ai pixels da analizzare,  $\underline{m}$  il centroide del cluster ed  $N$  il numero di componenti spettrali, si ha che le distanze possono essere scritte in questo modo:

- $L_1$  o city block:  $L_1(\underline{x}, \underline{m}) = \sum_{i=1}^N |x_i - m_i|$
- $L_2$  o Euclidea:  $L_2(\underline{x}, \underline{m}) = \|\underline{x} - \underline{m}\| = \sqrt{\sum_{i=1}^N (x_i^2 - m_i^2)}$

Le combinazioni di suddivisione dello spazio delle caratteristiche sono molteplici, il caso limite coincide in una situazione in cui ogni singolo pixel costituisce un cluster, ma ciò è privo di significato, è un caso teorico; la condizione più favorevole si ha quando i cluster sono molto compatti e ben distinti tra loro, questa è la condizione ideale.

Questa tipologia di classificazione è corredata di indicatori che esprimono le caratteristiche dei clusters:

- Matrice di dispersione totale  $S_{tot}$ : fornisce una misura dei vettori  $\mathbf{x}$  attorno alla loro media  $\mathbf{m}$  nello spazio multispettrale:

$$S_{tot} = \sum_i \sum_{x \in \omega_i} (\underline{x} - \underline{m})(\underline{x} - \underline{m})^T = N \cdot \underline{C}_{tot}$$

- Matrice di dispersione interna  $S_{int}$ : indica come sono dispersi gli  $n_i$  elementi dei vettori  $\mathbf{x}$  di ogni singolo cluster attorno alla propria

$$\text{media } \underline{m}_i: S_{int} = \sum_i \sum_{x \in \omega_i} (\underline{x} - \underline{m}_i)(\underline{x} - \underline{m}_i)^T$$

- Matrice di dispersione esterna  $S_{est}$ : indica il grado di separazione tra i diversi cluster:  $S_{est} = \sum_i n_i (\underline{m}_i - \underline{m})(\underline{m}_i - \underline{m})^T$

Inoltre vale il legame:  $S_{tot} = S_{int} + S_{est}$

Dove:  $\omega_i$  rappresenta un singolo cluster

$\underline{C}_{tot}$  è la matrice di covarianza totale

$$\underline{m} \text{ è la media totale: } \underline{m} = \frac{1}{N} \sum_i \sum_{x \in \omega_i} x$$

$$\underline{m}_i \text{ è la media del cluster } \omega_i: \underline{m}_i = \frac{1}{n_i} \sum_{x \in \omega_i} x$$

Per un set di osservazioni il raggruppamento dei pixel in cluster non è certo univoco, gli algoritmi di clustering operano cercando la suddivisione ottimale che è data da queste condizioni:

- $S_{int}$  minima, che significa cluster compatti;
- $S_{est}$  massima, a rappresentare cluster ben separati tra loro

Sapendo che la matrice di dispersione totale  $S_{tot}$  è costante l'algoritmo deve minimizzare  $S_{int}$  e massimizzare  $S_{est}$ . La minimizzazione avviene tramite la minimizzazione di una funzione scalare della matrice come può essere la traccia, in questo caso, minimizzare la traccia della matrice è equivalente a minimizzare la somma degli errori quadratici SSE (Sum of Squared Error).

$$SSE = \sum_i \sum_{x \in \omega_i} \|x - \underline{m}_i\|^2$$

Si procede valutando per ogni cluster la distanza media cumulativa di tutti i suoi elementi  $x$  dal baricentro  $\underline{m}_i$  del cluster, queste distanze vengono sommate per tutti i cluster: se la loro somma (SSE) è diminuita significa che le distanze degli elementi dal rispettivo centro del cluster si sono ridotte e quindi il cluster ottenuto è più compatto rispetto il precedente.

Si intuisce che questo procedimento deve essere eseguito per tutte le possibili partizioni di  $N$  in  $k$  clusters, si avranno un numero di possibilità quantificabile, ma non gestibile. Implementare l'algoritmo con questo criterio è molto complesso, per ovviare l'inconveniente, sono stati messi a punto algoritmi con approccio iterativo o di tipo gerarchico, capaci di ricavare soluzioni soddisfacenti. Gli approcci più diffusi sono quelli iterativi in cui viene richiesto a priori il numero di cluster da ricercare; negli approcci gerarchici, che fanno uso di alberi o grafi, invece il numero di cluster viene definito a posteriori.

### **Algoritmo delle medie mobili o *k-means***

Questo algoritmo viene generalmente utilizzato negli approcci di tipo iterativo. Concettualmente opera in questo modo:

- deve essere fissato il numero  $k$  dei clusters,
- la partizione dello spazio avviene in modo iterativo,
- l'elemento  $\mathbf{x}$  viene assegnato al cluster  $\omega_i$  che presenta la media  $m_i$  più vicina, la media dopo l'assegnazione viene ricalcolata e aggiornata,
- le iterazioni procedono finché non si arrestano le variazioni delle medie.

Operativamente devono essere svolti questi passaggi:

- vanno individuati  $k$  punti candidati a divenire centroidi  $\hat{m}_i$  ( $i=1, \dots, k$ ) dei clusters. La scelta dei centroidi è arbitraria ma, se vengono selezionati in modo distribuito sull'immagine, migliora la convergenza della procedura.
- Vengono prese in esame le posizioni dei pixels di training all'interno dello spazio multispettrale, secondo logiche di similarità il pixel viene assegnato al cluster che presenta il vettore delle medie più vicino. In questo modo si forma una prima separazione in classi dello spazio delle caratteristiche.
- A fine assegnazione vengono calcolati i nuovi vettori delle medie  $m_i$  ( $i=1, \dots, k$ ).
- Vengono controllati i nuovo baricentri, se nessun valore delle medie è cambiato, quindi si ha che  $\hat{m}_i = m_i$ , la procedura è terminata. Se ciò non accade allora si ripetono iterativamente i passi precedenti ponendo come valore  $\hat{m}_i^j = m_i^{j-1}$ , ovvero il valore del centroide al passo  $j$  è posto pari al valore della media del passo  $j-1$ .
- Il criterio di arresto del procedimento iterativo è dato dalla differenza dello spostamento delle medie da un'iterazione al passo  $j-1$  rispetto quella al passo  $j$ , è espresso da questa relazione:

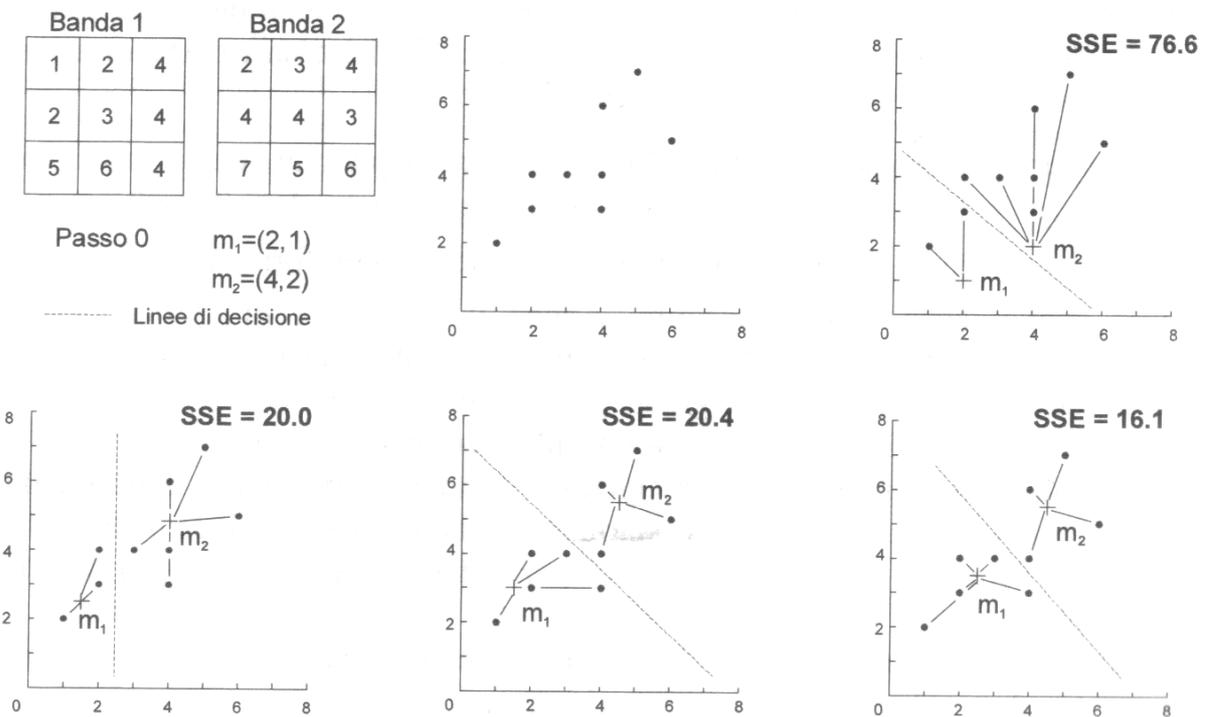
$$\Delta m(n) = \sum_{i=1}^k |m_i(n) - m_i(n-1)|$$

Questo procedimento garantisce una configurazione finale stabile che non dipende dalla posizione dei vettori delle medie scelte all'inizio. Ottenuti i vettori

delle medie della configurazione finale si può procedere alla classificazione tramite i criteri di distanza o col metodo di massima verosimiglianza.

Un ulteriore miglioramento della tecnica di clustering si ha con l'algoritmo *isodata*. Questo algoritmo esamina i clusters ottenuti: elimina quelli che sono formati da così pochi elementi perché privi di significato; accorpa quelli che sono tra loro molto vicini (merge); separa cluster che presentano anomale forme allungate (splitting) utilizzando criteri basati sulla deviazione standard delle bande spettrali.

La classificazione non supervisionata viene utilizzata in prevalenza quando non si hanno dei dati di training affidabili o per individuare in modo automatico il numero di classi rappresentative da ricercare in un'analisi supervisionata.



In figura sono mostrati alcuni esempi di clustering su due gruppi basati sul principio del vettore medio più vicino. In particolare devono essere notati lo spostamento delle medie tra iterazioni successive e la diminuzione progressiva della somma dei quadrati d'errore SSE.

### 2.2 Valutazione dell'accuratezza della classificazione

Le immagini classificate devono essere sottoposte ad una rigorosa valutazione statistica di accuratezza prima di essere consegnate all'utente utilizzatore. L'accuratezza di una mappa tematica è definita come la misura della fedeltà tra mappa tematica ottenuta dalla classificazione e un riferimento che viene assunto come verità a terra. Il concetto di accuratezza non è legato alla precisione della classificazione, con questo termine viene indicato il livello di dettaglio della classificazione. La conoscenza dell'accuratezza di una mappa deve essere sempre fornita per dare informazioni sull'affidabilità dell'elaborato e per indicare gli eventuali errori che sono presenti.

#### Errori ed accuratezza della mappa tematica

L'errore di classificazione viene commesso quando un pixel o un object, durante il processo di classificazione, viene assegnato ad una classe alla quale non appartiene. L'errore viene caratterizzato per frequenza, natura, grandezza e sorgente. La frequenza indica quante volte si ha l'errore di misclassificazione, la natura specifica le classi interessate, la grandezza esprime la gravità dell'errore, la sorgente indica la provenienza dell'errore.

La valutazione dell'accuratezza è eseguita mediante il confronto dell'immagine classificata e una seconda immagine che viene assunta come verità a terra, prodotta generalmente da informazioni più dettagliate e accurate. Queste immagini per essere confrontate devono essere compatibili rispetto alla scala, al livello di dettaglio, numero e tipologia di classi.

Esistono due metodologie per valutare l'accuratezza: *non-site-specific* e *site-specific*. Il primo metodo fa una valutazione confrontando i valori delle aree assegnate alle classi nelle due immagini; è un metodo elementare che non considera la localizzazione spaziale delle classi e questo limite può causare gravi errori di valutazione. L'accuratezza *site-specific* opera confrontando individualmente pixel per pixel delle due immagini, è un procedimento oneroso

ma garantisce una stima dell'accuratezza affidabile in quanto è considerata la localizzazione degli errori. Idealmente questo procedimento andrebbe applicato sull'intera immagine ma nella realtà dei fatti è complesso, quindi è applicato su alcune aree campione dette *test samples* che non devono assolutamente essere coincidenti ai training sites, perché significherebbe valutare la classificazione sulle informazioni utilizzate a produrla.

**La matrice d'errore**

La matrice d'errore contiene tutte le informazioni ricavate del confronto dei pixels dell'immagine classificata rispetto quella di riferimento; è una matrice quadrata di dimensione pari al numero delle classi. Lungo le colonne si hanno i dati del riferimento, nelle righe sono riportati i risultati della classificazione; gli elementi sulla diagonale principale riportano gli elementi correttamente classificati, gli elementi che sono al di fuori sono invece errori di classificazione. Nella figura sottostante è riportata la matrice d'errore per un'immagine che presenta  $q$  classi. Il termine  $p_{ij}$  rappresenta i pixel assegnati alla categoria  $i$  che in realtà sarebbero della categoria  $j$ .

		dati di riferimento (colonne)				Totali per classe nella mappa	
		classi	1	2	...	q	
dati di classificazione (righe)	1		$p_{11}$	$p_{12}$	...	$p_{1q}$	$p_{1+}$
	2		$p_{21}$	$p_{22}$	...	$p_{2q}$	$p_{2+}$
	⋮		⋮	⋮	...	⋮	⋮
	q		$p_{q1}$	$p_{q2}$	...	$p_{qq}$	$p_{q+}$
Totali per classe nel riferimento			$p_{+1}$	$p_{+2}$	...	$p_{+q}$	

La classificazione ideale è rappresentata da una matrice d'errore che non presenta elementi fuori dalla diagonale. La matrice d'errore oltre a fornire indicazioni riguardo la corretta classificazione, riporta dettagliate informazione sugli errori e sulla loro distribuzione.

### Misure d'accuratezza

La lettura della matrice d'errore consente di ricavare parametri, espressi in termini di probabilità, che esprimono la qualità della classificazione.

- Accuratezza Globale:  $AG = \frac{\sum_{k=1}^q p_{kk}}{n}$

dove con  $p_{kk}$  è indicata la somma degli elementi sulla diagonale e  $n$  in numero totale di test samples. AG rappresenta la probabilità che ha un punto estratto a sorte di essere correttamente classificato.

- Accuratezza Utente (User Accuracy):  $AU = \frac{p_{ii}}{p_{i+}}$  relativo alla classe  $i$

dove con  $p_{ii}$  è indicata la somma dei pixel assegnati correttamente alla classe  $i$ ,  $p_{i+}$  indica il numero totale dei pixel assegnati alla classe  $i$  dal processo di classificazione. AU rappresenta la probabilità condizionale che ha un punto estratto a sorte, appartenente alla classe  $i$ , di essere stato correttamente classificato. È un indicatore utile all'utilizzatore.

- Accuratezza Produttore (Producer Accuracy):  $PU = \frac{p_{jj}}{p_{+j}}$  relativo alla classe  $j$

dove con  $p_{jj}$  è indicata la somma dei pixel assegnati correttamente alla classe  $j$ ,  $p_{+j}$  indica il numero totale dei pixel della classe  $j$  nell'immagine di riferimento. PU rappresenta la probabilità condizionale che ha un punto estratto a sorte, appartenente alla classe  $j$  nell'immagine di riferimento, di essere stato correttamente classificato nella mappa tematica. È un indicatore utile al classificatore.

- Errore di Commissione:  $CE = 1 - AU = \frac{\sum_{k=1}^q p_{ik}}{p_{i+}}$  relativo alla classe  $i$  e ( $k \neq i$ )

È associato a AU; CE rappresenta la probabilità condizionale che ha un pixel estratto a sorte, appartenente alla classe  $i$  nella mappa classificata, di non essere stato correttamente classificato.

- Errore di Omissione:  $OE = 1 - AP = \frac{\sum_{k=1}^q p_{kj}}{p_{+j}}$  relativo alla classe  $j$  e ( $k \neq j$ )

È associato a AP; OE rappresenta la probabilità condizionale che ha un pixel estratto a sorte, appartenente alla classe  $j$  nella mappa di riferimento, di non essere stato correttamente classificato nella mappa tematica.

- Coefficiente di accordo kappa:  $\hat{k} = \frac{n \sum_{k=1}^q p_{kk} - \sum_{k=1}^q (p_{k+} \cdot p_{+k})}{n^2 - \sum_{k=1}^q (p_{k+} \cdot p_{+k})}$

Fornisce una stima dell'accuratezza di una mappa tematica depurata dall'effetto legato alle possibili coincidenze casuali tra le due mappe. Considera gli elementi della diagonale principale della matrice d'errore e indirettamente anche quelli che ricadono al di fuori. Il coefficiente K assume valori compresi tra 0 e 1 che rappresenta il massimo accordo tra i dati di riferimento e i dati generati dalla classificazione. Valori di K prossimi allo 0.4 indicano una classificazione scadente, valori intorno a 0.75 sono indice di una buona classificazione



# Capitolo III

## Estrazione di Features

Nel precedente capitolo sono state illustrati in modo generico gli algoritmi di classificazione. Nell'analisi delle immagini digitali il procedimento di classificazione riveste un'importanza fondamentale per l'estrazione di informazioni utili dai dati grezzi. I vari metodi che vengono utilizzati possono essere divisi in due categorie: metodi spettrali, o *pixel-based*, e metodi contestuali, o *object-based*.

L'evoluzione della tecnologia ha consentito di ottenere immagini ad alta risoluzione e ciò ha incentivato la ricerca di metodologie più moderne di classificazione in grado di sfruttare al meglio questa tipologia di dati. La novità consiste proprio nella metodologia di approccio al problema; la classificazione tradizionalmente opera su singoli pixel individuali, un approccio più moderno consiste nell'operare considerando le interrelazioni tra pixel contigui, l'approccio non è più dunque Pixel-Based ma diventa Object-Based. Quest'ultimo metodo consente di avere un maggior numero di features a disposizione nella fase di classificazione, che permettono di descrivere meglio e in modo più dettagliato gli oggetti di interesse secondo una metodica analoga all'interpretazione visiva fatta dall'uomo; è proprio questa caratteristica che lo rende più efficiente rispetto l'interpretazione eseguita con la metodologia pixel-based che ricorre alle sole informazioni spettrali. Per *features*, *attributi* o *caratteristiche* si intendono qui le proprietà che ha l'elemento di riferimento, pixel o object a seconda del metodo di classificazione, nello spazio delle caratteristiche; generalmente tra tutte quelle a

disposizione soltanto alcune sono in grado di rendere possibile l'individuazione dell'oggetto desiderato.

Le tradizionali procedure di classificazione pixel-based analizzano le informazioni spettrali di ogni singolo pixel; questo modo di operare è idoneo se applicato ad immagini a bassa risoluzione ma può risultare problematico se applicate su immagini ad alta risoluzione perché può creare un effetto sale e pepe.

La procedura di classificazione *object-based* non è più basata sui singoli pixel ma su aggregazioni di pixel che vengono chiamati *objects*. I più moderni programmi specialistici impiegati per lo studio di immagini digitali sono dotati di strumenti in grado di frammentare l'immagine in porzioni omogenee di pixel che presentano *features* simili. Questa operazione, che consente di ottenere gli *objects*, viene chiamata *segmentazione*.

Ormai molteplici studi sono stati sviluppati per accertare la validità e accuratezza di questa nuova procedura; Whiteside e Ahmad (2005) hanno dimostrato che, nel realizzare la mappa di copertura del suolo del Lichfiel National Park situato nel nord dell'Australia, la classificazione pixel-based funzionava bene in aree omogenee, ma la classificazione object-oriented forniva risultati migliori nell'identificazione di aree eterogenee. Il loro studio non ha interessato soltanto territori naturali ma anche un'area urbana dell'Australia. Anche le ricerche di Syed e al. (2005) su di un'area di studio urbana che presenta zone vegetate e spazi aperti hanno evidenziato la maggior flessibilità e miglior accuratezza della classificazione object-based rispetto quella tradizionale.

Gli studi di Robert C. Weih Jr. e Norman D. Riggan Jr. sono finalizzati al confronto tra queste due metodologie, ed in particolare a verificare se la procedura di classificazione object-based di immagini satellitari a media risoluzione e di immagini digitali aeree ad alta risoluzione produce una mappa di uso e copertura del suolo più accurata rispetto ad una classificazione supervisionata tramite una tradizionale metodologia pixel-based. Entrando più in dettaglio l'area di studio è ampia 16,850 ettari e riguarda Hot Springs, Garland Country in Arkansas: sono presenti un parco nazionale che presenta varie colture arboree da identificare,

corpi d'acqua, alcune cave nella roccia e la città di Hot Springs. Le immagini di questa area sono state acquisite dal satellite SPOT-5 in due istanti temporali distinti (inverno e primavera) e presentano due angoli d'incidenza diversi. Sono disponibili quattro bande: verde e rosso con risoluzione spaziale 10 metri, infrarosso vicino e infrarosso medio con risoluzione spaziale 20 metri. Oltre alle immagini satellitari SPOT-5 sono a disposizione un'immagine aerea true color RGB con risoluzione spaziale di 1 piede e un'ortofoto digitale ad infrarosso con risoluzione spaziale 1 metro. Tutte le immagini sono state ortorettificate tramite l'uso di Ground Control Points ed è stato applicato un ricampionamento per creare un unico dataset composto da 11 bande con risoluzione 1 metro. La classificazione object-based è stata ottenuta col software *Feature Analyst* che tramite algoritmi analizza il contenuto spaziale e spettrale degli objects che vengono soppesati in relazione ai loro vicini diversamente dalla tecnica pixel-based. Per entrambe le metodologie sono stati individuati dei training data per ogni classe da individuare per istruire il programma. L'accuratezza dei risultati è riportata tramite la matrice di confusione dove viene messa a confronto la classificazione automatica di aree di controllo, dette *check sites*, rispetto la conoscenza dell'effettiva copertura a terra. Nel caso di studio sono stati ottenuti i risultati riportati in tabella:

	OBJECT-BASED CLASSIFICATION	PIXEL-BASED CLASSIFICATION
<i>overall accuracy</i>	82.0 %	66.9 %
<i>user accuracy</i>	50.0 ÷ 100.0 %	22.2 ÷ 100.0 %
<i>producer accuracy</i>	50.0 ÷ 91.2 %	29.7 ÷ 94.7 %

I risultati di questo studio dimostrano che l'accuratezza complessiva del metodo di classificazione object-based è superiore se comparata con quella che si ottiene col metodo pixel-based supervisionato sulla stessa immagine.

Anche la disponibilità di immagini ad alta risoluzione di aree urbane, che risultano di difficile elaborazione spaziale e spettrale a causa della tipica eterogeneità che rende complessa l'analisi, ha incentivato lo studio della metodologia di classificazione object-based. Infatti la trattazione di immagini con questa risoluzione spaziale con i metodi tradizionali non porta risultati soddisfacenti anche per il basso numero di bande disponibili (4 per i sensori *Ikonos* e *QuikBird*) ed inoltre è bene correggere la distorsione geometrica e le variazioni radiometriche che si hanno tra due immagini per rendere il dataset spettralmente più omogeneo. Le informazioni più importanti per ricavare una carta dell'uso del suolo e una mappatura socioeconomica molto dettagliata sono ricavate dalla tessitura e dal contesto dell'immagine (Barnsley e Al. 1993).

Proprio su questo tipo di dati Herold, Scepan, Müller e Günther hanno condotto uno studio su 7 immagini multispettrali Ikonos con risoluzione spaziale di 4 metri a terra che interessano l'area urbana di Santa Barbara. Le immagini interessano differenti zone e sono state acquisite in istanti temporali distinti, quindi sono state necessarie una correzione atmosferica, una calibrazione radiometrica, per rendere omogenee le informazioni spettrali e una correzione geometrica per unire le immagini in una unica. I training sites sono identificati su un'immagine aerea. Il data set è stato analizzato con il software eCognition, consente una classificazione object-based; è stata fatta una segmentazione iniziale che viene personalizzata dall'utente impostando i parametri: *scale color shape smoothness compactness*, che vanno ad incidere sulla forma e dimensioni. La fase di classificazione è stata eseguita con una descrizione delle classi e assegnazione secondo un algoritmo di tipo nearest neighbor col criterio di minima distanza tra le features spaziali, spettrali, tessitura e forma. La difficoltà nel classificare con precisione aree urbane è concentrata nella discriminazione tra i vari manufatti: edifici, parcheggi, strade, ecc. perché presentano firme spettrali simili.

Questi ricercatori hanno deciso di operare con una struttura gerarchica basata su tre livelli; il primo è il più generico e comprende le classi: corpi d'acqua, edifici, superfici vegetate e superfici non vegetate, il secondo e il terzo sono

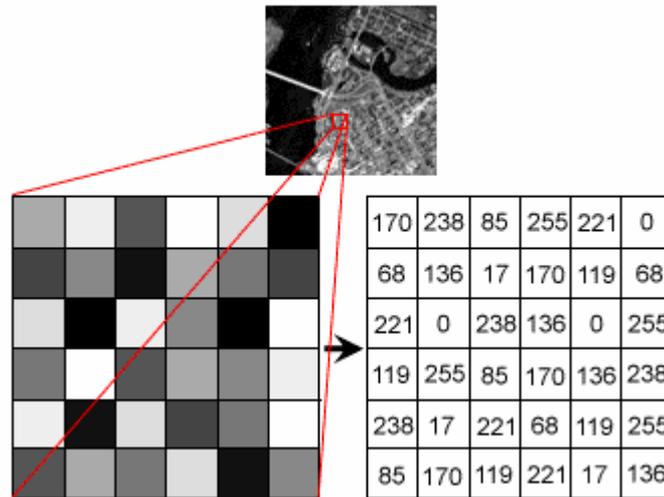
formati dalle sottoclassi, in particolare nel secondo livello le classi sono differenziate per le informazioni spettrali, il terzo livello, invece, fa uso anche delle informazioni di forma utili ad esempio per discriminare in modo corretto la classe delle strade da quella degli edifici.

Il risultato finale è stato testato con una valutazione dell'accuratezza, la matrice d'errore mostra come la classificazione risulti molto accurata per la vegetazione verde, i corpi d'acqua naturali e artificiali; buoni risultati sono stati ottenuti con alcune classi dell'urbano distinguibili grazie l'NDVI; i risultati peggiori si hanno per le classi strade e tetti scuri che vengono tra loro confuse. Ciò sottolinea come le informazioni aggiuntive che caratterizzano gli objects rispetto ai singoli pixels, come ad esempio la forma, siano molto utili in quei casi in cui due classi distinte presentino caratteristiche spettrali molto simili. L'accuratezza complessiva della classificazione risulta del 79%, un buon risultato se considerato il fatto che lo studio è stato effettuato su aree urbane che presentano un'elevata eterogeneità. Anche in queste circostanze la metodologia object-based ha prodotto buoni risultati.

### **3.1 La classificazione pixel-based**

L'approccio di classificazione pixel-based utilizza solo le informazioni di riflettanza contenute nei pixels dell'immagine. L'assegnazione delle classi avviene a livello del singolo pixel e dipende esclusivamente dal suo contenuto informativo: le uniche informazioni disponibili sono quelle relative al contenuto spettrale dell'immagine. Il problema con questo genere di classificazione è che viene utilizzata solo l'informazione spettrale, tralasciando completamente l'informazione fornita dai rapporti geo-spaziali all'interno dell'immagine. Gli algoritmi a disposizione per la classificazione di immagini digitali sono per la maggior parte di tipo convenzionale (Massima Verosimiglianza, Spectral Angle Mapper...) descritti nel capitolo precedente.

Con l'immagine sottostante si vuole chiarire il modo di operare di questa metodologia di classificazione: l'immagine viene scissa in pixels che vengono classificati uno ad uno in funzione dell'informazione radiometrica (digital number) con l'algoritmo adottato.



Questa metodologia di classificazione opera nelle condizioni ottimali se le classi da estrarre sono ben distinte nello spazio delle caratteristiche.

### 3.2 La classificazione object-based

La creazione di questa metodologia è motivata dal crescente interesse di poter disporre di strumenti capaci di fornire una migliore descrizione semantica del contenuto delle immagini. In molti casi le informazioni semantiche significative per il caso di studio, qualunque questo sia, non sono isolate nel singolo pixel, ma in genere si riscontrano in un *image-object*.

Gli *image-objects* sono raggruppamenti di pixels che presentano caratteristiche simili (luminosità, colore, tessitura) e vengono creati in un processo chiamato segmentazione. Sono di particolare interesse nel caso di immagine ad alta risoluzione dove uno stesso oggetto è rappresentato da una grande quantità di

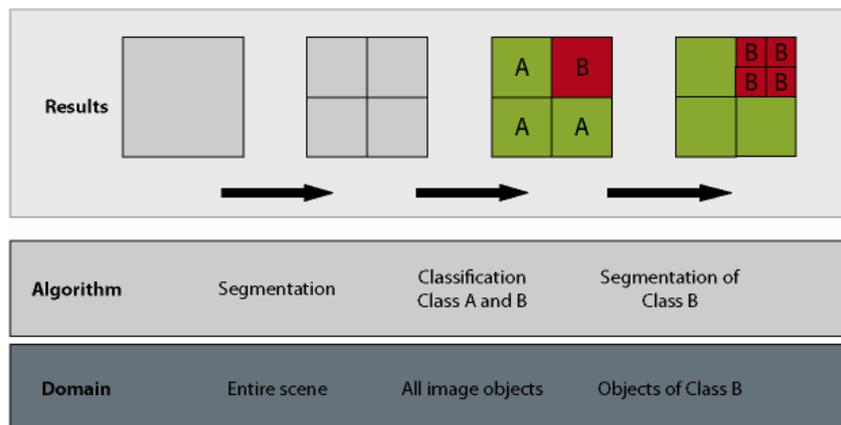
pixels, con questo assemblamento viene ridotto il numero di elementi da processare nella fase di classificazione.

Le informazioni spettrali contenute in un object derivano dai valori medi radiometrici dei pixels che ne fanno parte. Questo nuovo modo di operare utilizza anche un altro genere di informazioni aggiuntive, o features: forma, tessitura, relazioni geo-spaziali, contesto, rapporti dimensionali...

Una disponibilità di un numero maggiore di informazioni consente una migliore identificazione degli oggetti reali che possono essere individuati secondo criteri geometrici, relazioni tra image objects, tessitura, oltre alle caratteristiche spettrali, con una analogia al ragionamento d'interpretazione visiva da parte di un operatore umano.

Con questa tecnica di classificazione le unità di base dell'analisi sono gli image-objects; partendo dal presupposto che è di interesse estrarre dall'immagine informazioni sul mondo reale, è auspicabile una coincidenza degli oggetti reali con gli image-objects, cosa improbabile nella metodologia pixel-based.

Il processo di classificazione object-oriented è caratterizzato da due azioni: la *segmentazione* e la *classificazione*. Come mostrato in figura è obbligatorio iniziare con la segmentazione per creare gli image-objects, solo dopo sono possibili la classificazione e successive ripetizioni di queste azioni per identificare al meglio gli oggetti d'interesse.



Processo di classificazione object-oriented

### 3.2.1 La segmentazione

La segmentazione viene definita da Mather come “la ricerca di regioni omogenee in un’immagine col fine di classificarle”.

La segmentazione è un processo di suddivisione di un’immagine in porzioni separate (segments) che sono dei raggruppamenti di pixels vicini che presentano analoghe caratteristiche come lucentezza, colore, tessitura. I segments corrispondono idealmente ad oggetti del mondo reale. Sono state ideate svariate procedure per la segmentazione di immagini, tuttavia, soltanto alcune portano qualitativamente a risultati convincenti, robusti e con condizioni di funzionamento applicabili. L'aspettativa di un risultato di segmentazione è quello di estrarre automaticamente tutti gli oggetti di interesse contenuti nell’immagine; è un problema complesso a causa della difficile gestione della moltitudine dei contenuti semantici. Si potrebbe pensare di sviluppare algoritmi specifici applicabili solo a una ridotta tipologia di classi e di immagini. Si può generalizzare affermando che tipicamente si distinguono due fasi all’interno del processo di segmentazione:

- 1 una prima fase di modellazione delle caratteristiche estratte dalla tessitura dell’immagine e delle informazioni spettrali;
- 2 una seconda fase di ottimizzazione in cui le features sono raggruppate in segmenti omogenei seguendo un criterio di minimo costo di clustering.

Le tecniche di segmentazione possono essere suddivise nelle categorie: supervisionata e non supervisionata. La prima generalmente non viene impiegata in Telerilevamento perché non porta risultati ottimali. La segmentazione non supervisionata a sua volta è suddivisa in tre tipologie: point-based, edge-based, region-based, ma principalmente viene impiegata l’ultima.

Gli approcci più semplici per la segmentazione sono quelli a soglia globale che in genere portano a risultati di qualità ridotta. Gli algoritmi region-grow raggruppano i pixels partendo da un numero limitato di seed points, il punto baricentrico del cluster. Questi algoritmi dipendono fondamentalmente

dall'insieme dei seed points e difettano sul criterio d'arresto del processo d'assegnazione.

Gli algoritmo di segmentazione del tipo edge-based richiedono un solo parametro di input (scale level). Il livello di segmentazione è in funzione dello scale level; se viene impostato un basso valore verranno definiti pochi segments, viceversa un elevato valore dello scale level produce molti segments. Una buona segmentazione fa sì che la classificazione risulti più accurata quindi è bene individuare lo scale level maggiore che meglio delinea i bordi delle features che verranno cercate. Se il valore scelto dello scale level è eccessivamente alto, i confini dei segments non saranno giustamente definiti e si perderanno features d'interesse (segmentazione grossolana). La segmentazione, tramite la creazione degli objects, riduce notevolmente il numero di elementi da analizzare (nella tecnica pixel-based coincidono con i pixel).

In alcuni casi è possibile impostare alcuni parametri dell'algoritmo per dare un peso maggiore all'informazioni spettrali o alla tessitura, oppure alla forma e le dimensioni degli objects. Le dimensioni degli objects sono molto influenti perché incidono sul grado di dettaglio dell'analisi: un fattore di scala basso comporta poco dettaglio, al contrario un fattore elevato comporta una frammentazione marcata. L'ideale sarebbe identificare un valore del parametro *scale* che non comporta né sovra-frammentazioni né sotto-frammentazioni ma che riesca ad identificare i bordi degli oggetti d'interesse.

La formazione degli objects produce nuove informazioni in quanto durante la fase di classificazione saranno a disposizione features che riguardano le interrelazioni tra objects, tra neighbor-objects sia sullo stesso livello che tra livelli, si vengono a creare strutture ad albero che analizzano i contenuti informativi dell'immagine a diverse scale di dettaglio e sotto diversi aspetti.

Un prerequisito necessario per un'elaborazione d'immagini col metodo object-oriented è una corretta segmentazione dell'immagine. A causa delle diverse tipologie di analisi e di immagini disponibili le dimensioni degli objects devono essere adattabili alla scala d'interesse. Ciò è possibile soltanto tramite un

algoritmo di segmentazione generale basato su definizioni di omogeneità combinato a tecniche di ottimizzazione globale e locale che vengono complessivamente racchiuse all'interno del parametro *scala* che controlla le dimensioni medie degli image objects. Come regola grossolana si può considerare che gli image objects debbano essere di dimensioni superiori della scala del rumore e della tessitura dell'immagine.

### 3.2.2 Segmentazione multirisoluzione

Gli objects che si ottengono dalla segmentazione sono un dato grezzo che viene raffinato nelle procedure successive, si evidenziano di seguito le caratteristiche che devono presentare:

*alta qualità degli image object primitives*: gli object-primitives dovrebbero presentare una qualità universale applicabile e adattabile a molti problemi e anche a immagini arbitrarie della stessa tipologia;

*multirisoluzione*: l'estrazione di oggetti significativi dell'immagine deve tener conto della dimensioni del problema da risolvere; pertanto, la scala degli image-objects dovrebbe essere un parametro adattabile secondo le necessità del caso;

*risoluzione degli attributi dipendente dalla scala*: quasi tutti gli attributi degli image objects sono scala dipendenti, quindi possono essere comparati soltanto elementi che presentano dimensioni paragonabili;

*riproducibilità*: i risultati devono essere riproducibili per tutti gli utenti;

*universalità*: la segmentazione deve essere applicabile ad arbitrati tipi di dati e problemi;

*velocità di elaborazione*: le prestazioni devono essere ragionevoli anche su insiemi di dati di grandi dimensioni.

La qualità della classificazione è in funzione della qualità della segmentazione, perciò è determinante la scelta del software e dell'algoritmo che si andranno ad utilizzare. Per valutare la qualità della segmentazione è necessario definire dei criteri, sia qualitativi che quantitativi; di seguito si citano alcuni risultati degli studi di Baatz e Schape riguardo l'ottimizzazione dei risultati ricavati dalla segmentazione di immagini. [Baatz & Shape, 2000].

Stabilendo a priori l'eterogeneità e le dimensioni medie che devono avere gli image objects, hanno stabilito questi due possibili criteri quantitativi per la valutazione dei risultati derivati dalle segmentazione:

- 1 l'eterogeneità media dell'object deve essere minimizzata;
- 2 l'eterogeneità dell'image object, pesata sulle dimensioni in pixels, deve essere minimizzata.

Viene preferito il secondo criterio perché rende paragonabili objects di diverse dimensioni; significa in sostanza che l'eterogeneità media dei pixels deve essere minimizzata, ogni pixel viene pesato in relazione all'eterogeneità dell'image-object al quale appartiene.

I criteri qualitativi sono stati dedotti attribuendo importanza alla valutazione dell'occhio umano dell'operatore: all'interno dell'immagine enfatizzare i contrasti locali, segmentazione in objects di dimensioni più o meno simili, segmentare in funzione delle features che si vogliono ricavare.

La segmentazione multirisoluzione è un algoritmo di tipo bottom-up region-merging. L'algoritmo parte considerando ogni pixel un image object iniziale, procede aggregando pixels (*merge*) ad ogni passo secondo stabiliti criteri, andando a formare object di dimensioni maggiori. Il criterio decisionale per applicare il merge è basato sull'omogeneità locale che descrive la similarità tra image objects adiacenti, il processo viene arrestato quando l'incremento dell'omogeneità tra gli image objects supera il costo del merge. Al merge è assegnata una funzione di costo che esprime il grado di aggregazione dei pixels (*degree of fitting*); un valore elevato consente di effettuare più volte il merge producendo image objects di

grandi dimensioni, al contrario un valore basso genera object di piccole dimensioni. Per l'influenza sulle dimensioni degli objects, il *degree of fitting* viene comunemente chiamato parametro *scala* e viene definito dall'utente.

La segmentazione multirisoluzione è caratterizzata da due aspetti fondamentali:

- decisioni euristiche per determinare gli image objects che subiranno il merge;
- definizione di un concetto di omogeneità degli image objects per valutare ad ogni step il *degree of fitting* relativo a due image objects candidati al merge.

#### **Decisioni Euristiche**

Fissato il criterio di omogeneità è necessario fissarne uno ulteriore per individuare gli image objects candidati al merge, vengono usate le decisioni euristiche che forniscono un metodo di approccio alla soluzione di casistiche che presentano incertezze, al fine di generare nuova conoscenza. L'euristica di una teoria dovrebbe garantire uno sviluppo empirico tale da prevedere fatti nuovi non noti al momento dell'elaborazione della teoria.

Preso un image object arbitrario A, possono essere applicate diverse decisioni euristiche per trovare un image object B adiacente per il merge. Le seguenti quattro le possibilità sono caratterizzate da un incremento di vincolamento in termini di libertà di scelta:

- 1 **fitting:** il merge tra A e B avviene con la sola soddisfazione del criterio di omogeneità;
- 2 **best fitting:** il merge con A avviene per l'object B che soddisfa al meglio il criterio di omogeneità e presenta il minor valore del parametro scala;
- 3 **local mutual best fitting:** il merge tra A e B avviene se il criterio d'omogeneità soddisfatto è il migliore. Inoltre per B si ricerca il suo vicino C con il quale soddisfa al meglio il criterio di omogeneità. Ora si hanno due

possibilità o C coincide con A, oppure in caso contrario, si ripete lo stesso ciclo, ponendo B per A e C per B. Questo tipo di decisione euristica permette di trovare la coppia migliore di objects vicini ad A seguendo il criterio di omogeneità;

- 4 **global mutual best fitting:** il merge di A e B adiacenti avviene sull'intera scena e deve soddisfare al meglio il criterio di omogeneità. In questo caso l'ordine del procedimento è implicitamente data dalla euristica.

Tranne l'ultimo caso, le decisioni euristiche necessitano della definizione dell'object A per avviare il procedimento; inoltre il procedimento deve essere eseguito per più objects contemporaneamente per mantenere sotto controllo l'evoluzione delle dimensioni. Gli *starting points* possono essere definiti singolarmente per ogni ciclo, oppure si possono effettuare merge contemporanei sull'immagine ininfluenti fra loro.

### **Degree of fitting**

Prendendo in considerazione due image objects appartenenti allo stesso feature space  $d$ -dimensionale, il *degree of fitting* esprime la loro similarità, quindi la loro distanza nel feature space, ed è definito da questa relazione:  $h = \sqrt{\sum_d (f_{1d} - f_{2d})^2}$

La distanza tra i due objects può essere normalizzata e viene espressa nel seguente modo:

$$h = \sqrt{\sum_d \left( \frac{f_{1d} - f_{2d}}{\sigma_{fd}} \right)^2}$$

L'esecuzione del merge comporta un incremento dell'eterogeneità media di tutti gli image objects; l'obiettivo della procedura di ottimizzazione deve essere quello di minimizzare l'incremento di eterogeneità che inevitabilmente si presenta ad ogni singolo merge. L'eterogeneità spettrale di un image object può essere

definita dalla variazione dei valori medi spettrali o dalla deviazione standard dei valori medi spettrali. Il *degree of fitting*, di due image objects adiacenti, può essere allora meglio espresso dalla variazione di eterogeneità  $h_{diff}$  causata dal merge. Nasce la necessità di esprimere il cambiamento di eterogeneità prima e dopo un ipotetico merge di due objects vicini, alcune relazioni possibili sono:

- $$h_{diff} = h_m - \frac{h_1 + h_2}{2}$$

- $$h_{diff} = h_m - \frac{h_1 n_1 + h_2 n_2}{n_1 + n_2}$$
 dove  $n$  è la dimensione degli objects

- pesando l'eterogeneità in funzione della dimensione degli objects:

$$h_{diff} = (n_1 + n_2)h_m - (h_1 n_1 + h_2 n_2) = n_1(h_m - h_1) + n_2(h_m - h_2)$$

- generalizzazione della precedente considerando un numero di canali pari a  $c$  indicando con  $w$  il loro peso:

$$h_{diff} = \sum_c w_c [n_1(h_{mc} - h_{1c}) + n_2(h_{mc} - h_{2c})]$$

Occorre definire anche il concetto di eterogeneità di forma degli image objects. Un primo criterio può essere quello di paragonare il discostamento dalla forma ideale di oggetto compatto, espresso dal rapporto tra il perimetro  $l$  e l'area  $n$

in pixels: 
$$h = \frac{l}{\sqrt{n}}$$

Un secondo metodo è dato dal rapporto tra la lunghezza del perimetro  $l$  dell'object e la lunghezza del perimetro  $b$  di un rettangolo di selezione: 
$$h = \frac{l}{b}$$

I criteri di omogeneità di forma devono essere necessariamente affiancati ad un criterio di eterogeneità spettrale per avere significato; risultano particolarmente appropriati per immagini che presentano tessiture complesse con objects di forma irregolare o ramificata.

L'applicazione della segmentazione è strettamente legata al software utilizzato, nel seguente capitolo verranno illustrati il software Envi 4.5 e Definiens Developer che sono stati adottati nei casi di studio che riguardano questa tesi.



# Capitolo IV

## Software utilizzati

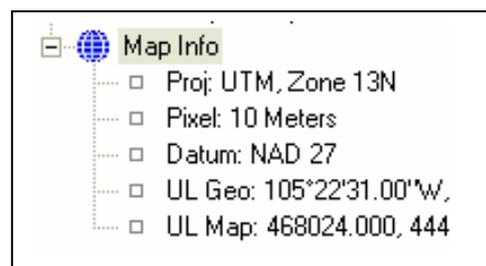
Sono stati utilizzati due differenti programmi che eseguono analisi di immagini iperspettrali ad alta risoluzione: Envi 4.5 e Definiens Developer 7.0.

### 4.1 Envi 4.5

ENVI è un software usato per il telerilevamento che opera su immagini satellitari in pancromatico, multispettrali, iperspettrali e radar consentendo la loro visualizzazione ed elaborazione a livelli di grigio, a colori o in falso colore; inoltre può lavorare sia con files in formato raster che vettoriale.

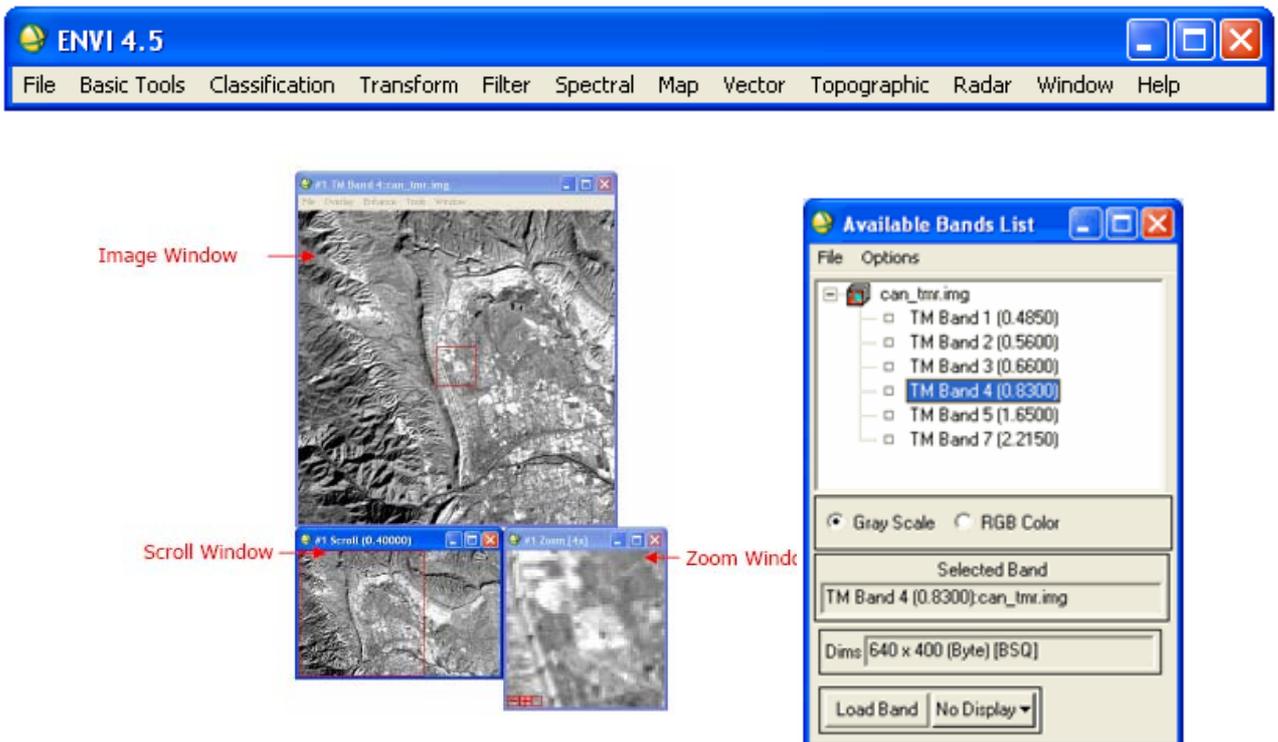
ENVI consente la visualizzazione di svariati files tra i quali `xxx.evf` formato proprietario del programma, `xxx.shp` shapefile da Arcview e ArcGis, `xxx.dxf` formato di Autocad. Il formato standard delle immagini gestito da ENVI è costituito da files a codifica binaria ai quali è associato un HEADER con codifica ASCII che presenta un'estensione del tipo `.hdr` con lo stesso nome dell'immagine.

L'interfaccia grafica del programma consiste in una barra menù principale, la finestra per la gestione dei files, dove per ogni file aperto compaiono le bande disponibili e la loro lunghezza d'onda caratteristica, e tre finestre che mostrano l'immagine in diversa scala: Image Window 1:1 (modificabile); Scroll Window in dimensioni ridotte tramite un fattore di



resampling; Zoom Window immagine ingrandita di un fattore 4x modificabile. MAP dINFO contiene le informazioni relative alla georeferenziazione ricavate dall'HEADER: la tipologia della proiezione, la dimensione del pixel a terra, il Datum, latitudine e longitudine geografiche e le coordinate della mappa.

Interfaccia grafica del programma:



Per ogni banda, tra le parentesi, viene indicata la lunghezza d'onda caratteristica. Visualizzando contemporaneamente diverse bande in finestre distinte è possibile relazionarle col comando link display: il puntatore del mouse si muoverà simultaneamente sulle finestre mostrando anche le coordinate del pixel se il file è georeferenziato. Consente la sovrapposizione multipla e dinamica per una facile comparazione delle immagini con possibilità di regolare la trasparenza.

**Funzionalità di ENVI 4.5:**

- STRETCHING: è un'operazione che modifica i Digital Numbers dell'immagine spalmando il loro valore lungo tutto l'intervallo da 0 a 255. Sono possibili diversi tipi di stretching: LINEARE, PIECEWISE LINEAR, GAUSSIANO, EQUALIZATION, SQUAREROOT. Inoltre, il comando INTERACTIVE STRETCHING permette all'utente di scegliere lui stesso come ridistribuire l'istogramma. Di default l'immagine viene resa più nitida applicando uno stretching del 2%.

- x, y, z, PROFILE - FIRME SPETTRALI: il programma può estrarre firme spettrali da dati iperspettrali con possibilità di confrontarle con librerie da letteratura. Z PROFILE (SPECTRUM) visualizza la firma spettrale del pixel selezionato. Le firme spettrali della libreria, rispetto a quelle ricavate dalle bande, non sempre hanno valori omologhi di riflettanza quindi per uniformarli c'è il comando EDIT(x,y) SPACE FACTOR e si può agire sui due assi.

- IMAGE TRANSECTS

- DENSITY SLICING

- SCATTEROGRAMMI i pixel delle immagini sono caratterizzati dai DN e dalla loro posizione.

- SCATTERPLOT 2D: in ascissa venne riportato il valore che quel pixel assume in una certa banda e in ordinata il valore relativo ad un'altra.

- SUBSETTING: Si usa il comando RESIZE DATA

- SPETTRALE: si crea una nuova immagine con un numero ridotto di bande rispetto l'originale

- SPAZIALE: si crea una nuova immagine che è una porzione di quella originale

- COLOR TABLE: serve per vedere una singola banda a colori associandole una palette. Il programma fornisce diverse possibilità. Alcune palettes associano colori chiari a DNs bassi per meglio evidenziare i bordi degli oggetti.

- MOSAIKING FILES: unisce due files
- CONTOUR LINES: Per DEM è in grado di generare le curve di livello in modo automatico. Viceversa note le curve di livello si può generare un DEM col comando TOPOGRAPHIC.
- MAPPE DI CLASSIFICAZIONE A COLORI
- DEFINIZIONE E PROCESSAMENTO DELLE ROI
- QUICK FILTER PREVIEW
- CREAZIONE E VISUALIZZAZIONE DI VETTORI

**Moduli di ENVI 4.5**

ENVI ZOOM	Permette la gestione di molteplici files contemporaneamente.
DEM EXTRACTION	Consente l'estrazione dei dati di elevazione da immagini PUSHBROOM STEREO ma anche ASTER e IKONOS.
FLAASH (fast line of sight atmospheric analysis of spectral hypercubes)	è un modulo per rimuovere gli effetti ottici di offuscamento causati dall'atmosfera.
ENVI FEATURE EXTRACTION	è un modulo di envi zoom che consente di estrarre informazioni da immagini pancromatiche ad alta risoluzione e da immagini multi spettrali basandosi su caratteristiche spaziali e spettrali. Consente di estrarre una grande varietà di features. Envi Feature Extraction è progettato per lavorare con ogni tipo di immagine fornitagli, magari già pre-elaborata, è in grado di impiegare meno tempo a comprendere i dettagli.
NITF/NSIF	Questo modulo supporta la lettura dei file NITF (National Imagery Transmission Format) e NSIF (Nato Secondary Image Format)
GEODATABASE SUPPORT	Consente la lettura di data-set raster e vettoriali creati con ARCGIS
S.I.D.  Spectral information divergence	È un nuovo metodo di classificazione spettrale che misura la divergenza tra due pixel collegati dallo spettro di riferimento. Più e piccola la divergenza più sono simili.
INDIPENDENT COMPONENT (IC) ANALYSIS	Consente di trasformare un set di segnali random mescolati tra loro in componenti che sono mutuamente indipendenti. IC può distinguere features anche quando si dispone di piccole porzioni di immagini.

### 4.1.1 Envi Zoom: estrazione delle features

Envi Feature Extraction è un modulo in grado di estrarre informazioni da immagini ad alta risoluzione spaziale e/o spettrale. Può essere usato per estrarre una gran varietà di features come veicoli, costruzioni, strade, ponti, fiumi, laghi. Una feature è un fenomeno geografico definito dall'utente che può essere modellato o rappresentato tramite l'uso di un data-set geografico. Questo programma consente di lavorare con qualunque tipo di immagine e ha lo scopo di diminuire il tempo necessario per interpretare i risultati.

Uno degli aspetti innovativi del Feature Extraction è che è in grado di mostrare in anteprima i risultati attraverso il Preview Portal. Envi Feature Extraction utilizza un approccio OBJECT BASED per classificare la immagini e ciò è un vantaggio mettendolo a confronto con le tecniche tradizionali PIXEL BASED (l'immagine viene classificata usando le informazioni spettrali contenute nel pixel). Si deduce che quest'ultimo modo di classificazione può lavorare meglio con dati iperspettrali ma non è idoneo per il trattamento di immagini pancromatiche e multispettrali.

Il metodo OBJECT BASED offre una maggior flessibilità nell'estrazione di features su immagini ad alta risoluzione in pancromatico e multispettrali perché classifica un set di object invece che classificare i pixels individualmente. Per OBJECT viene intesa una regione di interesse spaziale, spettrale (luminosità e colore) e/o caratteristiche della tessitura. Inoltre consente di estrarre features multiple allo stesso tempo.

Il preprocessing dell'immagine è possibile con ENVI prima di operare col Feature Extraction.

Eventuali elaborazioni dell'immagine che possono agevolare la classificazione:

**SPATIAL DOWN SAMPLING:** Riduce la risoluzione spaziale dell'immagine di input per facilitare il caricamento. Se l'immagine di input è molto

pesante è bene ricampionarla per velocizzare il processamento oppure eseguire un subsetting spaziale.

ENVI PREPROCESSING TOOLS: Per meglio individuare le features è opportuno ridurre il numero delle bande conservando soltanto con quelle più idonee.

LAYER STACKING: Può essere usato per costruire un nuovo file multibanda formato dall'immagine di input arricchita da altre informazioni come DEM, immagini Lidar, shapefiles. Questi attributi extra possono considerevolmente migliorare e correggere il risultato della classificazione.

SPATIAL FILTERING: Se l'immagine di input contiene molto rumore si possono applicare dei filtri spaziali come il LEE, idoneo a preservare i confini, e quello della mediana.

#### FEATURE EXTRACTION WORKFLOW

L'architettura del programma per l'estrazione delle features è riportata nello schema nella pagina seguente. Il WORKFLOW è organizzato in steps che non sono rigidi, ma che consentono di ritornare sui passi precedenti cambiando impostazione senza dover ripartire daccapo.

Envi feature extraction guida l'utente lungo il processo dell'immagine ed è scomposto in due steps:

FIND OBJECT: ha inizio con la segmentazione dell'immagine in regioni di pixel, formazione degli OBJECTS e calcolo degli attributi. Prevede quattro passi:

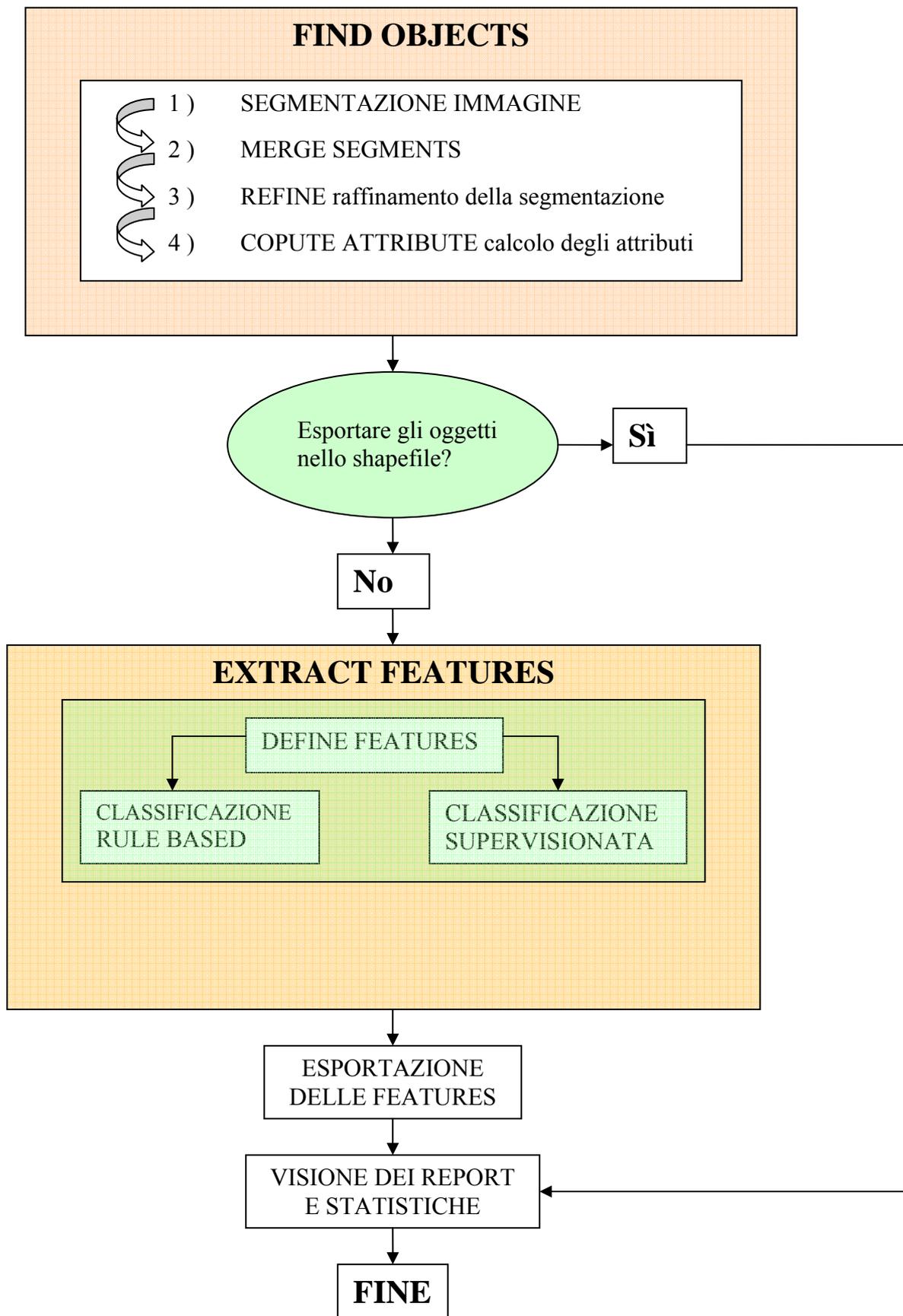
SEGMENT

MERGE

REFINE

COMPUTE ATTRIBUTE

EXTRACT FEATURES: determinazione delle features e classificazione dell'immagine.



I seguenti comandi sono attivi ad ogni passo:

Previous: consente di ritornare al passo precedente per modificare le impostazioni;

Next: accetta il setting e procede al passo successivo;

Preview: visualizza in tempo reale l'anteprima dei risultati nel preview portal.

Terminata la prima fase si può personalizzare l'estrazione delle features procedendo con la classificazione che può essere supervisionata o rule based. Infine i risultati vengono esportati in uno shapefile e/o in un'immagine raster.

## 1. FINDING OBJECTS

### SEGMENTAZIONE DELL'IMMAGINE

La segmentazione è un processo di suddivisione di un'immagine in porzioni separate (segments) che sono dei raggruppamenti di pixels vicini che presentano analoghe caratteristiche come lucentezza, colore, tessitura. I segments corrispondono idealmente ad oggetti del mondo reale.

Envi zoom impiega un algoritmo di segmentazione del tipo "EDGE-BASED" che richiede un solo parametro di input (scale level) ed è molto rapido. Il livello di segmentazione è in funzione dello scale level; se viene impostato un basso valore verranno definiti pochi segments, viceversa un elevato valore dello scale level produce molti segments. Una buona segmentazione fa sì che la classificazione risulti più accurata quindi è bene individuare lo scale level maggiore che meglio delinea i bordi delle features che verranno cercate. Se il valore scelto dello scale level è eccessivamente alto i confini dei segments non saranno giustamente definiti e si perderanno features d'interesse (segmentazione grossolana). Se lo scale level è pari a 100,0 si ottiene un solo segment. È bene assicurarsi che più di una feature d'interesse non venga raggruppata nello stesso segment insieme ad altre e, viceversa, è altresì opportuno che la stessa feature non venga troppo frammentata.

Questa ultima situazione viene detta over-segmentation e deriva da un basso valore di scale level. Se questi problemi sono inevitabili durante la segmentazione possono essere ridotti nella seconda fase. Durante questa prima fase non è possibile agire sugli istogrammi dell'immagine, sono disponibili invece funzioni come lo zoom, il pan, funzioni per visualizzare layers sovrapposti: trasparenza, blend, flicker, swipe e il preview portal che mostra in tempo reale i risultati.

Quando si è soddisfatti dei risultati si prosegue cliccando next per procedere alla fase di merge.

Terminato il processo di segmentazione Envi ZOOM produce un'immagine detta Region Means che consiste in un file raster dove ogni segments è caratterizzato dal valore medio delle bande dei pixels che lo compongono.

### MERGE DELL'IMMAGINE

Questo step è facoltativo e serve per aggregare piccoli segments che fanno parte di vaste aree come boschi, nuvole, terreni, dove la sovra segmentazione può essere un problema. A titolo di esempio, se l'immagine mostra una fitta vegetazione, se si incrementa lo scale level della segmentazione, si rischia che non vengano sufficientemente delineate le aree prive di copertura arborea. (problema della sotto-segmentazione). Anche in questo passo è in funzione il Preview Portal per mostrare l'anteprima dell'applicazione. Più è elevato il livello di merge più i segments verranno diminuiti in numero. Il scala del merge level rappresenta la soglia del parametro  $\lambda$  che può assumere valori da 0.0 a 100.0 (impostata a 100.0 i segments vengono mescolati e riportati ad uno solo). Envi Zoom utilizza l'algoritmo Full Lambda Schedule ideato da Robinson, Redding e Crisp (2002). Questo algoritmo mescola iterativamente i segmenti adiacenti basandosi su di una combinazione spaziale e spettrale delle informazioni. L'algoritmo procede se trova un paio di regioni,  $i$  e  $j$ , tali che il costo di mergering  $t_{i,j}$  sia minore del valore predefinito della soglia del parametro  $\lambda$ .

Costo del mergering:

$$t_{i,j} = \frac{\frac{|O_i| \cdot |O_j|}{|O_i| + |O_j|} \cdot \|u_i - u_j\|^2}{length(\partial(O_i, O_j))}$$

Dove:

$O_i$  = è la regione i-esima dell'immagine;

$O_j$  = è la regione j-esima dell'immagine;

$|O_i|$  = è l'area della regione i-esima dell'immagine;

$|O_j|$  = è l'area della regione j-esima dell'immagine;

$u_i$  = è il valore medio della regione i-esima;

$u_j$  = è il valore medio della regione j-esima;

$\|u_i - u_j\|^2$  = è la distanza euclidea tra i valori spettrali delle regioni i-esima j-esima;

$length(\partial(O_i, O_j))$  = è la lunghezza del bordo in comune ad  $O_i$  e ad  $O_j$

Se fosse possibile, per ottimizzare i risultati, sarebbe più ragionevole applicare il merge localmente soltanto alle zone sovrasegmentate.

Terminata la fase di merge viene creata un'immagine in output chiamata "Merged Region Means" che verrà aggiunta nel layer manager e verrà rimossa l'immagine Region Means.

Se il mergering non è un'opzione di giovamento all'immagine lo si può evitare procedendo allo step di refine impostando il merge level al valore di default (0,0).

## REFINING SEGMENTS

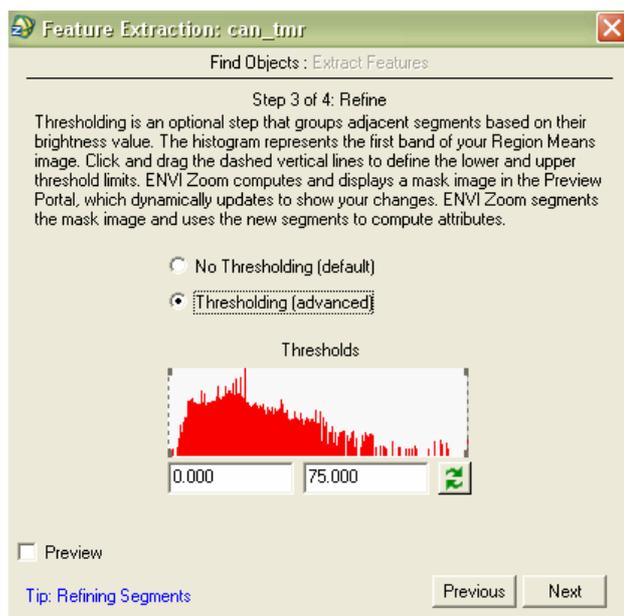
È possibile raffinare il risultato della segmentazione usando un altro metodo opzionale di mergering detto "Thresholding" utile per l'estrazione di features

puntuali come aeroplani, costruzioni, dove la sovrasedimentazione diventa un problema. Il thresholding è un'operazione raster che lavora sulla Region Means Image e raggruppa i segments vicini basandosi sul loro valore di luminosità (Digital Number). Funziona bene con features che sono in alto contrasto col loro sfondo.

La finestra del processo di estrazione delle features propone in default “no thresholding”, così facendo si proseguirà saltando questa opzione, in alternativa si ha il thresholding con le relative impostazioni.

Setting del thresholding:

le linee verticali dell'istogramma devono essere trascinate per definire i limiti inferiore e superiore dell'intervallo. Attivando l'opzione Preview in monitor sono



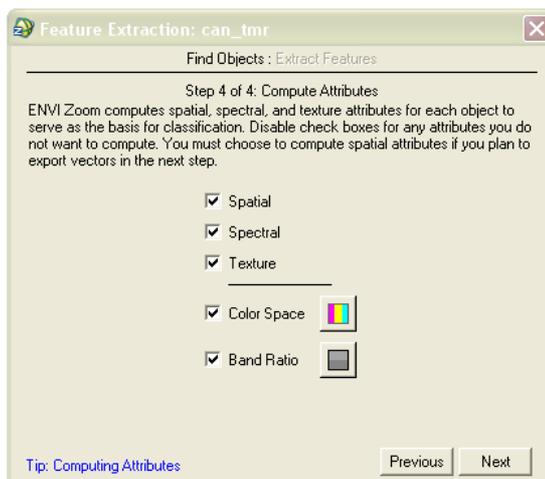
mostrati i cambiamenti in modo dinamico dell'immagine, a seguito della modifiche sull'istogramma, che possono essere comparati con l'immagine di partenza tramite l'uso del comando “trasparenza”. L'istogramma in ascissa ha il valore del DN, in ordinata, invece, la frequenza dei pixels che presentano quel valore. Muovendo le soglie sull'istogramma si

seleziona il range dei DN d'interesse ai quali viene assegnato un unico valore DN=255 (bianco); ai restanti esclusi dalle soglie verrà assegnato il valore DN=0 (nero). Quando si definiscono i limiti del Thresholding sostanzialmente si definiscono i nuovi contorni delle features d'interesse.

Per ogni nuova regione che si viene a formare viene associato un identificatore che verrà usato per il calcolo degli attributi. Terminato il processo verrà generata una Refine Region Means Image.

## COMPUTING ATTRIBUTES

Envi Feature Extraction utilizza la classificazione object based che offre il vantaggio di studiare gli oggetti per la loro varietà di attributi spaziali, spettrali e tessitura e scegliere quelli più rappresentativi. Il calcolo degli attributi viene influenzato dalle impostazioni (preferences) del programma e quindi è opportuno controllare i valori assunti dai parametri.



Tipologie di attributi che vengono calcolate sono:

### 1. SPAZIALI

AREA: è l'area totale del poligono esclusi i fori valutata in unità della carta/mappa/immagine

LENGTH: è il perimetro del poligono compreso i fori interni in map units

COMPACT: è la misura della compattezza del poligono e della sua forma:

$$COMPACT = \frac{\sqrt{(4 \cdot AREA) / \pi}}{LENGTH}$$

Il cerchio è la figura geometrica più compatta con valore di

$$COMPACT_{cerchio} = \frac{\sqrt{(4 \cdot \pi \cdot r^2) / \pi}}{2 \cdot \pi \cdot r} = \frac{1}{\pi}$$

Per un quadrato invece  $COMPACT_{quadrato} = \frac{\sqrt{(4 \cdot l^2)/\pi}}{4 \cdot l} = \frac{\sqrt{1/\pi}}{2} = \frac{1}{2\sqrt{\pi}}$

CONVEXITY: misura la convessità del poligono. Assume valori pari a 1 per poligoni convessi senza fori, per poligoni concavi assume valori minori a 1.  
CONVEXITY = length of convex hull / LENGTH

SOLIDITY: è una misura della forma del poligono. Compara l'area del poligono a quella di un'altra figura convessa che la circoscrive.

$$SOLIDITY = \frac{AREA}{AREA \text{ del poligono circoscritto}}$$

SOLIDITY=1 poligoni convessi senza fori:

SOLIDITY<1 poligoni convessi che presentano fori

ROUNDNESS: compara l'area del poligono a quella di un quadrato avente lato pari al massimo diametro del poligono. Il massimo diametro viene determinato come lunghezza dell'asse maggiore di un box, orientato e vincolato, che racchiude il poligono.

$$ROUNDNESS = \frac{4 \cdot AREA}{\pi \cdot MAXAXISLEN^2}$$

$$ROUNDNESS_{cerchio} = 1$$

$$ROUNDNESS_{quadrato} = \frac{4}{\pi}$$

MAXAXISLEN: è la lunghezza dell'asse maggiore di un box orientato e vincolato che racchiude il poligono. Il valore è in unità della mappa o in pixel. Se l'immagine non è georeferenziata sono riportate solo le unità in pixels

MINAXISLEN: è la lunghezza dell'asse minore di un box orientato e vincolato che racchiude il poligono.

FORMAFACTOR: rapporta l'area del poligono al quadrato del perimetro:

$$FORMAFACTOR = \frac{4 \cdot \pi \cdot AREA}{(perimetro_{tot})^2}$$

$$FORMAFAC\text{TOR}_{\text{cerchio}} = \frac{4 \cdot \pi \cdot \pi \cdot r^2}{(2 \cdot \pi \cdot r)^2} = 1$$

$$FORMAFAC\text{TOR}_{\text{quadrato}} = \frac{4 \cdot \pi \cdot l^2}{(4 \cdot l)^2} = \frac{\pi}{4}$$

ELONGATION: è il rapporto degli assi, maggiore e minore, del poligono calcolati col metodo del box. Assume valori pari ad 1 per quadrati e maggiori di 1 per rettangoli.

$$\text{ELONGATION} = \text{MAXAXISLEN} / \text{MINAXISLEN}$$

RECT\_FIT: è una misura che indica come la forma del poligono venga descritta più o meno bene da un rettangolo. Assume valori pari ad 1 per rettangoli e minori di 1 per le altre figure.

$$\text{RECT\_FIT} = \text{AREA} / (\text{MAXAXISLEN} \cdot \text{MINAXISLEN})$$

MAINDIR: è il valore in gradi dell'angolo sotteso dall'asse maggiore del poligono rispetto l'asse x. MAINDIR= 90° corrisponde alla direzione N/S; 0-180 corrisponde alla direzione E/O.

NUMHOLES: corrisponde al numero di fori che sono presenti all'interno del poligono.

HOLESOLRAT: è il rapporto tra l'area totale del poligono compresi i fori rispetto l'area totale senza i fori:  $\text{HOLESOLRAT} = \frac{\text{AREA} + \text{AREA}_{\text{fori}}}{\text{AREA}}$

Le formule per calcolare: COMPACT, CONVEXITY, SOLIDITY, ROUNDNESS e FORMAFAC\text{TOR}, sono state ideate da Russ (2001).

Envi Zoom calcola gli attributi spaziali non basandosi sulla geometria originale ma, su una versione "smussata" delle forme applicando l'algoritmo di Douglas e Peucker (1973).

Calcolare gli attributi sulla geometria elaborata dall'algoritmo garantisce che le misure di forma siano meno sensibili alla rotazione dell'oggetto.

### 2. SPETTRALI

I seguenti attributi sono calcolati per ogni banda  $x$  dell'immagine originale:

MINBAND\_ $x$ : è il minimo valore assunto dai pixels nella regione della banda  $x$

MAXBAND\_ $x$ : è il massimo valore assunto dai pixels nella regione della banda  $x$

AVGBAND\_ $x$ : è il valor medio assunto dai pixels nella regione della banda  $x$

STDBAND\_ $x$ : è il valor della deviazione standard assunto dai pixels nella regione della banda  $x$

### 3. TESSITURA

TX\_RANGE: data range medio dei pixels contenuti all' interno del kernel. Il kernel è un vettore di pixels utilizzato per racchiudere un'operazione di un subset di pixels.

TX\_MEAN: valor medio dei pixels contenuti all' interno del kernel.

TX\_VARIANCE: varianza media dei pixels contenuti all' interno del kernel

TX\_ENTROPY: valor medio di entropia dei pixels contenuti all' interno del kernel

Color Space e Band Ratio sono disponibili soltanto per immagini multispettrali con almeno 3 bande. Sono funzioni opzionali che consentono di calcolare questi attributi:

BAND RATIO: assume valori compresi tra -1 e 1.

HUE (tinta): viene spesso utilizzato come filtro colorato e viene misurato in gradi ( $0^\circ$ - $360^\circ$ ) con  $0^\circ$ =rosso,  $120^\circ$ =verde,  $240^\circ$ =blu;

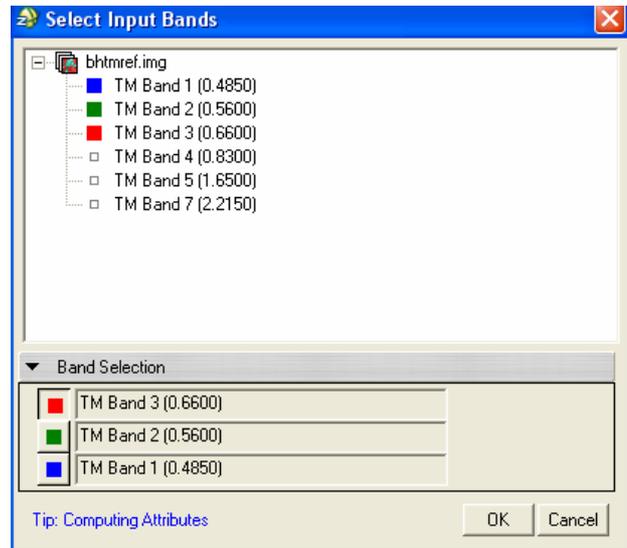
SATURATION: viene utilizzato come filtro colorato ed è misurato in valori di floating points da 0.0 a 1.0

INTENSITY: fornisce una misura migliore della luminosità rispetto all'attributo spettrale AVGBAND\_x. Viene misurato in valori di floating points da 0.0 a 1.0

### COLOR SPACE

Al fine di calcolare gli attributi HSI (tinta, saturazione, intensità) è necessario avere in input un file multispettrale con almeno 3 bande.

Se il file contiene le informazioni sulle lunghezze d'onda, allora Envi riesce autonomamente a calcolare gli attributi del color space; se non sono disponibili l'utente deve istruire il programma, altrimenti vengono utilizzate le prime tre bande.



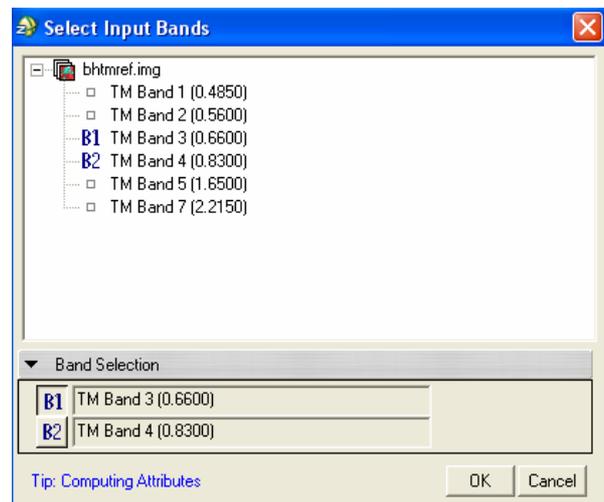
### BAND RATIO

Con questo comando Envi calcola una banda normalizzata ottenuta dal rapporto fra due bande usando l'equazione:

$$-1.0 < \frac{B_2 - B_1}{B_2 + B_1 + eps} < 1.0$$

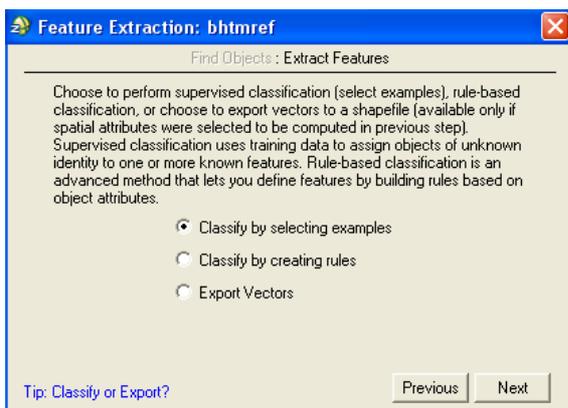
Dove: eps: minor valore necessario affinché non avvenga una divisione per 0

$B_1, B_2$  : bande utilizzate



Nel caso in cui  $B_1, B_2$  rappresentino rispettivamente le bande del rosso e dell'infrarosso si ottiene l'indice di vegetazione normalizzato (NDVI) che assume valori elevati per vegetazione verde.

## 2. ESTRAZIONE DELLE FEATURES



La finestra guida l'utente attraverso i processi di definizione delle features e di classificazione del tipo object based che può essere rule-based o supervisionata.

I risultati della classificazione si possono esportare in uno shape file vettoriale o in formato raster. Il

programma permette tre alternative delle quali può esserne scelta soltanto una. Sono:

### Exporting Vectors

Classify by selecting examples che porta ad una classificazione supervisionata

Classify by creating rules che porta ad una classificazione rule-based

### Exporting Vectors

Consente di esportare gli oggetti vettoriali e gli attributi degli object trovati senza definire le features e fare la classificazione.

La funzione export vectors è disponibile soltanto quando gli attributi spaziali vengono calcolati.

Abilitando la funzione Export Attributes gli attributi vengono salvati in un file XXX.dbf associato allo shape file di output.

Al termine del processo viene visualizzato il report dell'intero iter del feature extraction

## Classify by selecting examples (Classificazione supervisionata)

Questo processo utilizza i training data per assegnare agli objects con identità incognita una o più feature conosciute.



Maggiore è la quantità dei training samples selezionati maggiore è il miglioramento dei risultati della classificazione.

### DEFINIZIONE DEI TRAINING DATA

Il training data consiste in un oggetto che viene selezionato come rappresentativo per le feature conosciute.

Anche se sempre non sembra, alcune features, come mostrato nella figura a fianco, sono altamente segmentate quindi frazionate in molteplici objects con caratteristiche cromatiche differenti. Il comando show boundaries è utile nel caso citato perché mostra la sovra frammentazione non sempre percepibile ad occhio nudo. Affinché un object diventi un training data è sufficiente che venga selezionato. Per ottenere risultati migliori è bene scegliere per ogni distinta feature una varietà di objects che presentano diverse sfumature del colore che provengono da differenti zone.

Si possono gestire le features personalizzandole, ad esempio editando il loro nome, o rimuovendole.

I training data selezionati e i parametri della classificazione possono essere salvati in un file XXX.xml; in questo modo è possibile rendere disponibili i vecchi training data files per la classificazione corrente. Il salvataggio è possibile se sono presenti almeno features che contengono almeno un object ciascuna. Ciò consente di sperimentare come differenti impostazioni influenzano i risultati e confrontarli per poi determinare la soluzione migliore.

Per creare dei buoni training data è bene:

- provare più combinazioni per ottenerne una migliore;
- usare i training data provenienti da una prima immagine per classificare una seconda immagine simile;
- definire un unico training data file per più immagini ed usarlo per classificare altre immagini simili.

La seconda e la terza condizione sono opzioni avanzate che possono essere utili se si devono analizzare parecchie immagini dello stesso sensore che sono state prese sotto le stesse condizioni ambientali e si è interessati ad estrarre lo stesso set di features da ogni immagine. Il termine “similari” in questi casi si riferisce ad immagini catturate dallo stesso sensore e con analoghe condizioni ambientali. Facendo un esempio non si può utilizzare un training data file estratto da un’immagine quikbird per classificare un’immagine Landsat. Per ottenere migliori risultati le immagini dovrebbero possedere la stessa geometria di proiezione, lo stesso numero di bande e gli stessi parametri di lavoro (scale level, merge level, thresholds). Si possono salvare e recuperare i training data anche come alternativa a realizzarli per ogni immagine.

Un training data file contiene le seguenti informazioni:

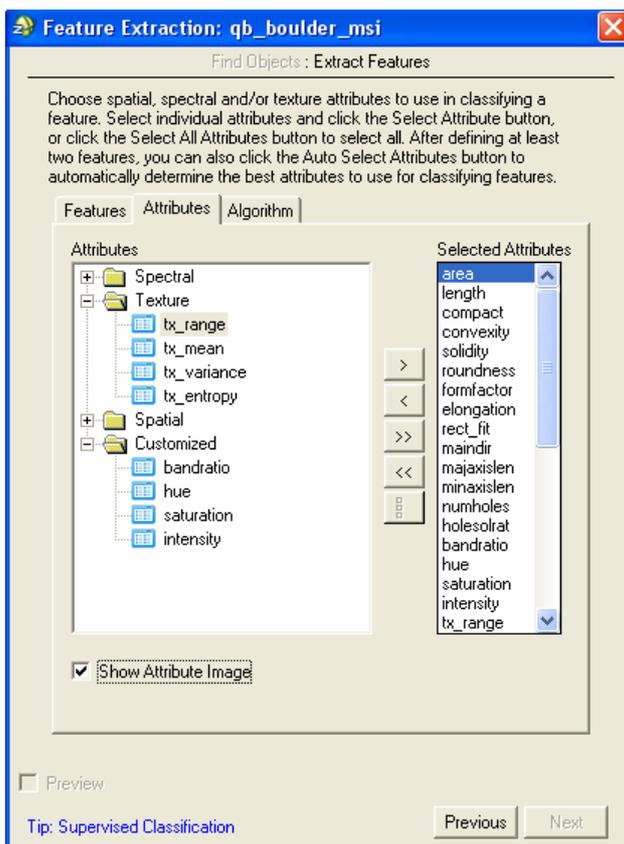
- i parametri utilizzati nel Feature Extraction workflow per generare gli objects (Scale Level, Merge Level, Refine parameters, e gli attributi calcolati);
- i metodi di classificazione e i parametri associati;
- gli attributi ed i training samples per una o più immagini

Per poter utilizzare un training data precedente, va ripristinato e per questo fine è necessario che coincidano i parametri impostati nel Feature Extraction workflow per generare gli objects. Non occorre invece che le due immagini abbiano la stessa dimensione spaziale però devono avere lo stesso numero di bande. Cioè non si può processare un’immagine multispettrale con un training data

file creato da immagini pancromatiche. Inoltre è possibile ripristinare soltanto un solo training data ad immagine.

Quando si lavora con un training data file ripristinato è possibile aggiungere o cancellare delle features dalla feature list.

Dopo aver definito almeno due features contenenti almeno un oggetto ciascuna (training sample) con la funzione preview si possono visualizzare i risultati della classificazione sul preview portal. Aggiungendo o rimuovendo features, oggetti selezionati o attributi il preview portal aggiorna dinamicamente la classificazione. È bene però sottolineare che l'opzione di preview rallenta le performance. In seguito si può procedere direttamente all'esportazione; è permesso creare fino a 47 features.



Al fine della classificazione è possibile scegliere tra attributi spaziali, spettrali e tessitura in quanto, in relazione al caso specifico, soltanto alcuni attributi sono più utili di altri a tal fine perchè sono in grado di distinguere meglio gli objects rendendo i risultati della classificazione più accurati. Quando sono selezionati tutti gli attributi quelli irrilevanti possono generare rumore nei risultati. Si ricorda che ENVI Zoom calcola gli attributi con l'unico scopo della classificazione e che

questi non vengono allegati in nessun shapefiles di output.

Gli attributi sono separati in quattro classi: Spectral, Texture, Spatial e Customized. Quest'ultima raggruppa il Color Space e il Band Ratio (attributi calcolabili soltanto per immagini multispettrali).

Per selezionare gli attributi si devono usare queste icone:

**Select Attribute:** Seleziona un solo attributo da portare nella lista degli attributi selezionati. Se si seleziona una cartella verranno selezionati tutti gli attributi che contiene.

**Unselect Attribute:** rimuove un solo attributo per volta dalla lista degli attributi selezionati.

**Select All Attributes:** con questo comando vengono portati nella lista tutti gli attributi.

**Unselect All Attributes:** con questo comando vengono rimossi dalla lista tutti gli attributi.

**Auto Select Attributes:** con questo comando il programma calcola gli attributi più idonei alla classificazione delle features. Questa funzione è basata sugli studi di Yang (2007). L'icona è attiva se sono presenti almeno 2 features che contengono 2 objects ciascuna. Questo comando è utile quando si ha un gran numero di objects.

Il risultati della classificazione non sono influenzati dall'ordine di scelta degli attributi.

È presente una sezione con opzioni avanzate per selezionare l'algoritmo di classificazione ed impostarne i parametri.

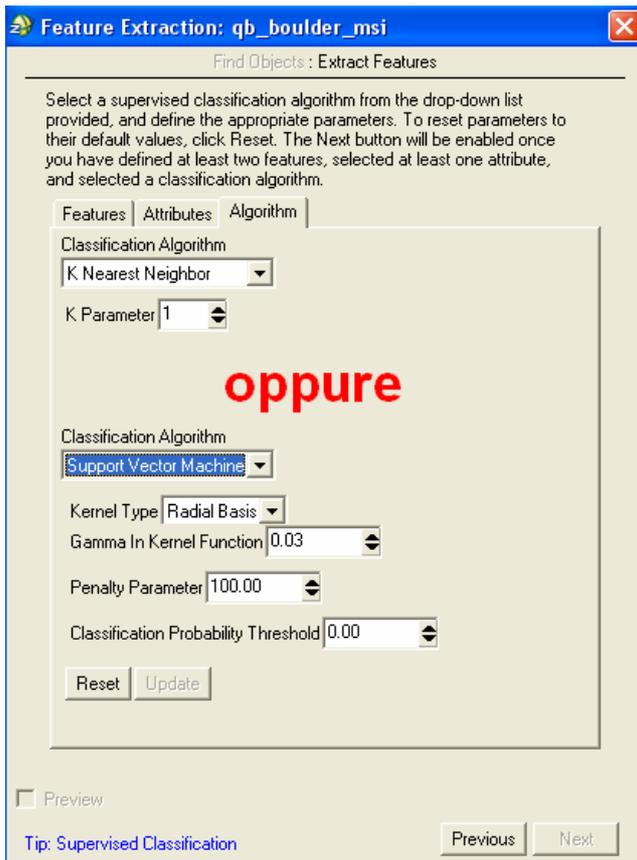
È possibile scegliere tra gli algoritmi K Nearest Neighbor e Support Vector Machine.

Anche in questo caso con l'opzione preview viene mostrato come variano i risultati modificando i parametri.

### Algoritmo K Nearest Neighbor

Il metodo di classificazione K Nearest Neighbor considera la distanza euclidea nello spazio n-dimensionale dei training data dove n è pari al numero di attributi degli objects selezionati per la classificazione. Questo metodo

generalmente è più robusto della tradizionale classificazione nearest-neighbor perché la K nearest distances è utilizzata come parametro principale per determinare a quale classe appartengono i target. Inoltre è meno suscettibile al rumore del dataset ed in generale produce dei risultati di classificazione più accurati se comparati ai tradizionali metodi nearest-neighbor.



Il parametro K è posto pari al numero di vicini considerati nel processo di classificazione. Il valore di default è 3; si può introdurre un numero intero dispari anche superiore a 3. è consigliabile sperimentare con svariati valori per ricercare il miglior risultato. Cambiando il valore del parametro i risultati cambiano considerevolmente. La scelta ideale del parametro K dipende dal dataset e dai training data selezionati. Valori elevati in genere riducono gli effetti del rumore però

facilmente portano a classificazioni inaccurate. Valori tipici sono: 3, 5 e 7.

Questo algoritmo assegna al target non la classe del neighbor più vicino ma, analizzando un numero di vicini pari parametro al K, quella che appartiene al numero maggiore di neighbors in quell'intorno.

### Support Vector Machine

Support Vector Machine (SVM) è un sistema di classificazione ottenuto da teorie statistiche. Questo metodo richiede la conoscenza di svariati parametri che a

loro volta dipendono dalla tipologia del kernel scelto. Si ricorda che il kernel è un vettore di pixels utilizzato per racchiudere un'operazione tra un subset di pixels.

L'algoritmo SVM consente l'uso di Quattro diversi tipi di kernel: Linear, Polynomial, Radial Basis Function e Sigmoid.

Il kernel del tipo Radial Basis Function è quello più idoneo nella maggior parte dei casi ed è proposto in default.

Ogni tipologia di kernel è correlata un diversa funzione matematica:

Linear:  $K(x_i, x_j) = x_i^T x_j$

Polynomial:  $K(x_i, x_j) = (x_i^T x_j + r)^d > 0$

Radial Basis Function (RBF):  $K(x_i, x_j) = \exp(-\|x_i - x_j\|^2) > 0$

Sigmoid:  $K(x_i, x_j) = \tanh(x_i^T x_j + r)$

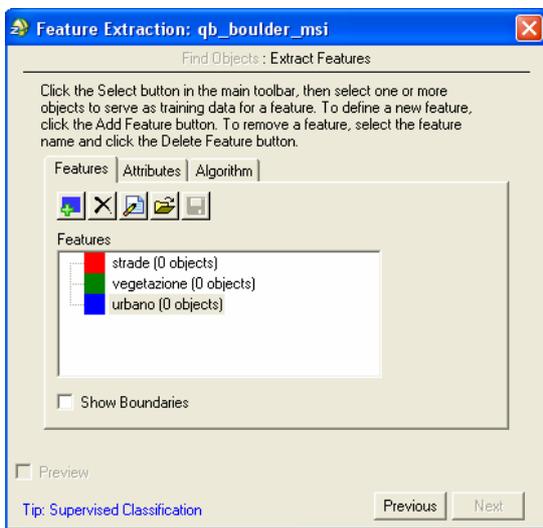
Il parametro Gamma in Kernel Function è un numero floating-point superiore a 0.01. In default è pari all'inverso del numero degli attributi calcolati. Questo parametro va definito per tutte le tipologie di kernel tranne quella "Linear". Il parametro Bias rappresenta il valore di r nel caso di kernel di tipo Polynomial o Sigmoid; in default è posto pari a 1.

Se il kernel è di tipo Polynomial, è necessario impostare il grado che assume: 1 come valore minimo (default) e 6 come massimo. Valori alti garantiscono una maggior accuratezza nella delineazione dei bordi tra le classi. Impostando 1 per questo parametro si ottiene una funzione polinomiale di primo grado: due classi distinte saranno separate da una retta. Questo caso è equivalente ad utilizzare il kernel di tipo Linear che però funziona bene se si hanno due classi ben distinte.

Il Penalty Parameter è un numero floating-point che assume valori superiori allo 0.01; il valore di default è 100.0. Questo parametro controlla l'entità dell'errore permesso e forza i margini tra le classi. Incrementare il valore del penalty parameter fa crescere il costo dei pixels malclassificati. In sostanza maggiore è il valore si ha una maggiore quantità di training data che può non essere considerate

Il parametro Classification Probability Threshold serve ad impostare il valore di probabilità necessaria all'algoritmo SVM per classificare i pixels. I pixels che presentano valori del parametro inferiore alla soglia non saranno classificati. Il valore di default è 0.0 e può essere portato fino a 1.0; valori elevati causano una maggior quantità di pixels inclassificati.

Per ripristinare i valori di default di tutti i parametri è sufficiente selezionare il pulsante Reset.



### **Classificazione Rule Based**

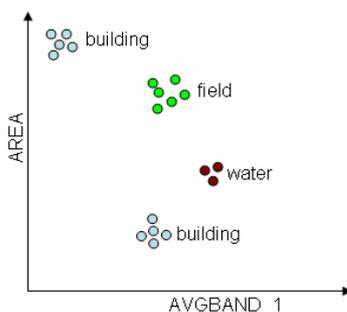
È una metodologia avanzata di classificazione che definisce le features tramite regole basate sugli attributi degli objects definite dall'utente.

La classificazione rule-based è un potente strumento per l'estrazione delle features e in genere fornisce risultati migliori rispetto la classificazione supervisionata per

certe features.

La costruzione delle regole è primariamente basata sulla conoscenza e sul ragionamento della persona a proposito di alcune specifiche features.

Per spiegare questo concetto si può pensare ad una strada che è allungata, gli edifici hanno una forma rettangolare, la vegetazione presenta un elevato valore per l'indice NDVI e gli alberi hanno una tessitura superiore rispetto un terreno.



Dopo aver assunto questi concetti il passo successivo consiste nel definire regole composte da una o più condizioni.

Ogni oggetto rappresenta un data point (sample) nello spazio n-dimensionale degli attributi.

In questo spazio i gruppi di sample dello stesso tipo vengono chiamati clusters. Si possono definire molte regole da dare alle features, in particolare in quei casi dove gli oggetti assegnati appartengono a più di un cluster.

Nasce la necessità di definire molteplici regole per quelle classi sparse nello spazio n-D. (come mostrato in figura classe building)

### Fuzzy Logic

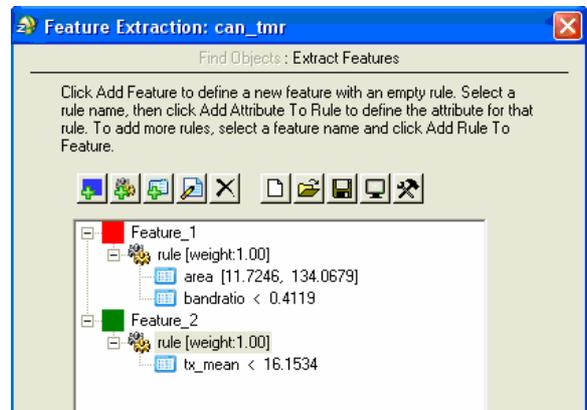
La classificazione tradizionale rule-based è basata su rigorose regole binarie, ovvero gli objects che non rispettano tutte le regole di appartenenza ad una feature non vengono assegnati e rimangono inclassificati anche se le rispettano al 90%.

La classificazione rule-based in ENVI Feature Extraction può essere di due tipi: un primo che osserva le rigide regole binarie, l'altro chiamato Fuzzy Logic. Quest'ultimo non utilizza le rigide regole binarie (tutto vero o tutto falso) ma sistemi di funzioni per rappresentare il grado o la percentuale di appartenenza dell'oggetto alle features.

L'estrazione d'informazioni da dati di remote sensing è limitata da diversi fattori quali: rumore del sensore, ridotta risoluzione spettrale e spaziale, degradazione del segnale, imprecise separazioni tra le classi di uso del suolo. Molte immagini di remote sensing contengono dei mixed pixels ovvero nello stesso pixel sono presenti più classi.

Il Fuzzy Logic aiuta ad alleviare questi problemi simulando l'incertezza o informazioni parziali che rende consistenti grazie al ragionamento umano.

L'output di ogni Fuzzy Rule è una mappa di confidenza dove i valori rappresentano l'attendibilità che l'oggetto appartenga ad un certo tipo di feature



definita dalle regole. Nella classificazione l'oggetto è assegnato alla feature che ha il massimo valore dell'intervallo di confidenza.

Con la classificazione Rule-Based si possono variare i gradi di Fuzzy Logic per ogni condizione mentre si costruiscono le regole.

Anche per questa tipologia di classificazione sono a disposizione tutti gli attributi suddivisi nelle categorie: Spectral, Texture, Spatial e Customized.

Ad ogni feature può essere associata più di una regola formata da più condizioni sugli attributi selezionati. Il peso della regola può variare tra 0 e 1(default). Se si imposta un valore inferiore a 1 viene assegnato alla regola un intervallo di confidenza più basso; ciò è utile per quelle regole che sono meno robuste rispetto le altre.

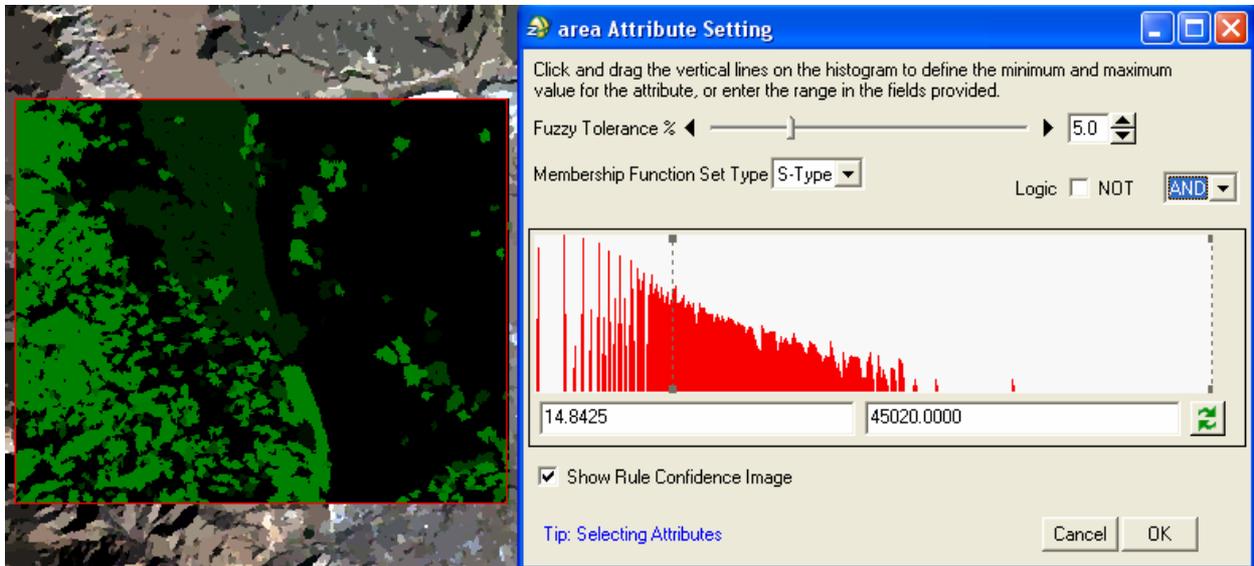
È possibile ripristinare rule set già esistenti o salvare quelli attuali in formato xxx.xml

Quando viene selezionato un attributo compare un'immagine in scala di grigi che rappresenta il valore assunto dall'attributo in ogni object della scena. Queste immagini aiutano a selezionare gli attributi più opportuni a definire le regole che caratterizzano la feature. Se gli attributi che appartengono alla feature hanno un alto contrasto in relazione agli altri oggetti significa che quegli attributi ben rappresentano la feature ed è giusto che appartengano alla regola. In questa fase può essere d'aiuto il comando della trasparenza dell'immagine per confrontare l'immagine degli attributi con quella iniziale. Cambiando attributo l'immagine si aggiorna.

Per creare la regola deve essere scelto un attributo; comparirà un istogramma che mostra la frequenza di avvenimento lungo l'asse y, in ascissa sono riportati i valori assunti dall'attributo per tutto l'insieme dei vector objects dell'immagine. Muovendo le linee verticali si definisce il range di valori che può assumere la feature.

L'opzione Show Rule Confidence Image è abilitata di default, modificando i limiti dell'istogramma nel Preview Portal viene aggiornata la probabilità degli

objects di appartenere a quella feature. Se un object viene rappresentato con un colore chiaro è alto il grado di appartenenza alla feature altrimenti se è di colore scuro fino al nero non appartiene.



Il Confidence Threshold in default è posto pari a 0.4 quindi se la luminosità di un object assume valori superiori a 102 verrà classificato secondo la regola che descrive la feature.

Il grado di fuzzy logic relativo alla regola può essere impostato muovendo la barra Fuzzy Tolerance per meglio modellare l'incertezza inerente alla classificazione. Il valore di default è del 5.0% incrementabile fino ad un massimo del 20.0 %.

Quando si imposta il grado di Fuzzy Tolerance su di un elevato valore, più objects saranno in una transizione di appartenenza/non appartenenza alla feature. Se si imposta il grado di Fuzzy Tolerance pari allo 0.0 % il fuzzy logic non verrà eseguito.

I gradi, relativi alla singola regola, di Fuzzy Logic vengono regolati sulla barra "fuzzy tolerance %" per meglio modellare l'incertezza inerente alla classificazione. Il valore di default è pari al 5%, quello massimo è 20%. Quando si lavora con regole multiple o con regole definite da attributi multipli è bene non impostare la fuzzy tolerance a zero.

Il fuzzy logic consente di meglio approssimare l'incertezza con un ragionamento umano così che l'oggetto abbia un'alta probabilità di appartenere alla classe assegnata.

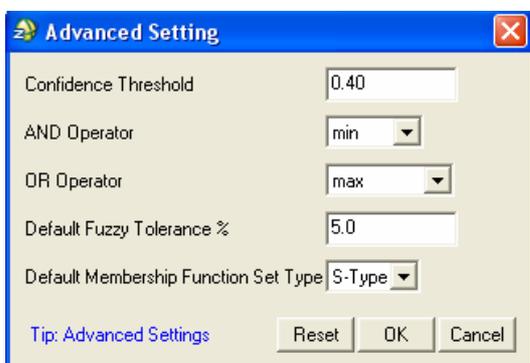
Le Membership functions sono concetti matematici per modellizzare la distribuzione dei dati.

Nelle finestra compare il menù a tendina Membership Function Set Type dove è possibile selezionare "S-Type" (default) oppure "Linear". ENVI Feature Extraction mette a disposizione queste due membership functions per modellizzare i concetti fuzzy come: "maggiore di" "minore di" e "tra". Le teorie di Jin e Paswaters (2007) sono utilizzate nelle impostazioni della S-Type e della Linear membership functions.

La scelta tra le due Membership Function dipende dalla distribuzione dei valori di ingresso delle variabili. Con gli stessi valori di Fuzzy Tolerance, la S-type membership function procura un minor grado di fuzziness rispetto alla Linear membership function.

Selezionando Logic NOT vengono applicati i valori esterni al range selezionato.

Per concludere la definizione delle impostazioni va trattato il menu a tendina: AND (default) oppure OR. Questa opzione è applicabile soltanto se per la stessa regola sono definiti almeno due attributi. AND e OR sono degli operatori logici per combinare la condizione corrente con quella precedente.



Il tasto  detto Display Rule Set mostra in modo ordinato le rules che definiscono le features.

Cliccando il tasto  detto Advanced Settings compare una finestra che mostra le impostazioni avanzate e consente di

personalizzarle. Durante la classificazione, l'oggetto non viene assegnato se il suo massimo valore di confidenza è inferiore al Confidence Threshold (posto a 0.4 di default).

Per gli operatori AND e OR sono possibili diverse opzioni che vengono applicate all'intero rule set.

Facendo un esempio:

supponiamo che la membership function per la condizione C1 valga  $m_1$  e che la membership function per la condizione C2 valga  $m_2$ .

Per l'operatore AND si ha:

min: (default) il valore di confidenza di C1 AND C2 è pari a:  $\min(m_1, m_2)$ .

product: il valore di confidenza di C1 AND C2 è pari a:  $m_1 * m_2$ .

Per l'operatore OR si ha: max: (default) il valore di confidenza di C1 OR C2 è pari a:  $\max(m_1, m_2)$ .

probabilistic: il valore di confidenza di C1 OR C2 è pari a:

$m_1 + m_2 - m_1 * m_2$ .

Per cambiare il valore di default del Fuzzy Tolerance, negli istogrammi interattivi, lo si può modificare in "Default Fuzzy Tolerance %". Questa opzione può essere utile quando si vuole utilizzare lo stesso valore per l'intero rule set. Cambiare questo valore non influenza le regole precedentemente esistenti.

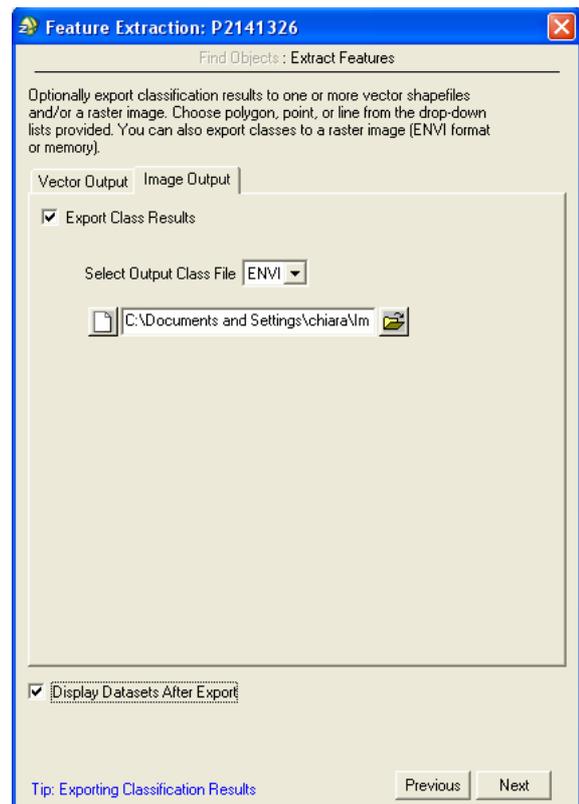
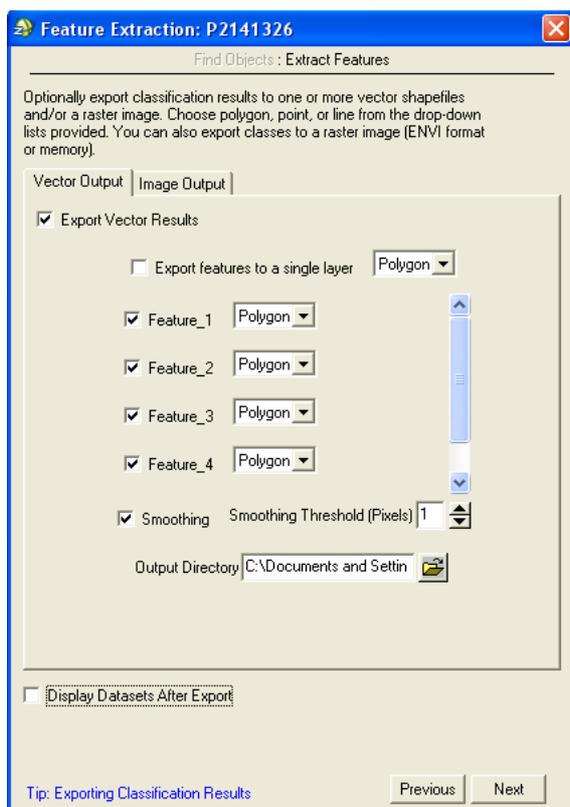
Default Membership Function Set Type consente di definire il tipo di membership function che verrà utilizzata di default negli istogrammi interattivi. Questa opzione può essere utile quando si vuole utilizzare la stessa membership function per l'intero rule set. Cambiare questa impostazione non influenza le regole precedentemente esistenti.

Dopo aver definito almeno una feature con una regola, cliccando preview dalla finestra di dialogo si ha l'anteprima della classificazione. La feature senza regole vengono ignorate.

## ESPORTARE I RISULTATI DELLA CLASSIFICAZIONE

Terminata la classificazione, Supervisionata o Rule Based, si procede al passo di Export.

Nella finestra sono presenti due schede



La seconda scheda, chiamata Image Output, consente di esportare i risultati della classificazione in un'immagine raster. Il file in output viene realizzato nel formato proprietario di ENVI. Quando si crea uno shapefile e/o un file immagine di output, si può scegliere di visualizzare subito i risultati selezionando "Display Datasets After Export".

La prima scheda, chiamata Vector Output, consente di esportare i risultati della classificazione in un file vettoriale. È possibile esportare i risultati della

classificazione in un unico shapefile o salvare ogni singola feature in uno shapefile distinto.

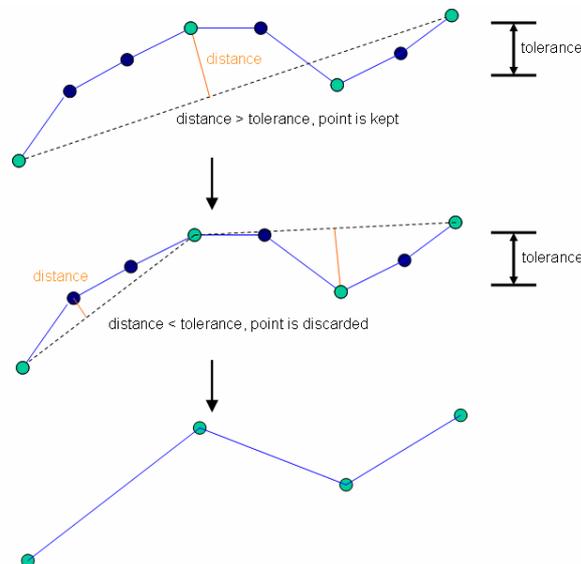
È possibile soltanto una di queste opzioni:

- esportare tutte le features in un singolo layer in un formato vettoriale del tipo: Polygon, Point, o Line. Questa opzione di default è disabilitata.
- esportarle in shapefiles distinti ogni singola features. Anche in questo caso va selezionato per ogni feature: Polygon, Point, o Line dal menu a tendina.

Gli objects sono dei poligoni; selezionando Line vengono trasformati in centreline invece, selezionando Point, vengono trasformati in uno shapefile formato soltanto da punti che rappresentano i centri dei poligoni che rappresentano gli objects. Gli shapefiles che verranno creati saranno chiamati con lo stesso nome della feature che rappresentano.

### Smoothing Vectors

È un'opzione che usa l'algoritmo di Douglas-Peucker. Questo algoritmo utilizza il minimo numero di vertici per descrivere una feature in una data scala. È più idoneo a generalizzare features come fiumi e non objects come edifici. L'algoritmo inizia individuando il primo e l'ultimo punto della polilinea o del poligono che descrive l'object e traccia una linea immaginaria che li congiunge come viene mostrato nella figura sottostante.

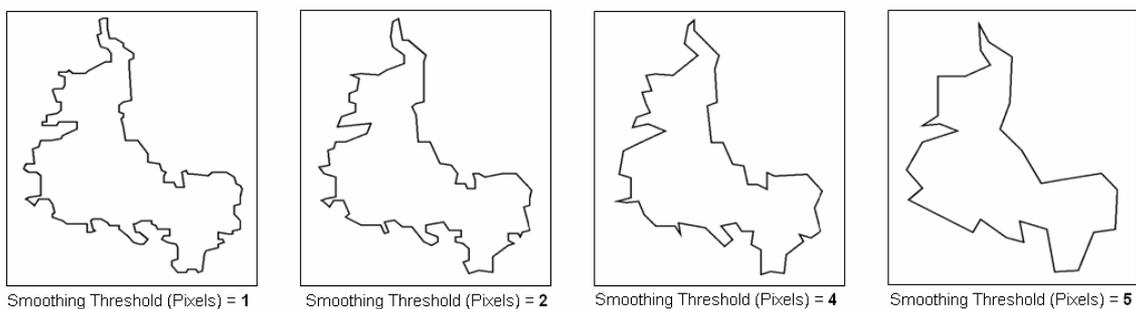


Per tutti gli altri punti compresi tra il primo e l'ultimo, l'algoritmo identifica quelli che hanno una distanza ortogonale dalla linea congiungente superiore alla tolleranza. Il valore della tolleranza viene introdotto in Smoothing Threshold ed è pari al numero di pixels utilizzato per generalizzare una polilinea o un poligono vettoriale..

Se sono presenti punti che distano oltre la tolleranza, l'algoritmo prende questi punti e crea tante polilinee quanti sono i punti più uno. L'algoritmo ripete il processo con ogni nuova polilinea finché nessun punto disterà dalla polilinea più del Smoothing Threshold. I punti che vengono conservati durante l'ultima interazione descriveranno la nuova feature, gli altri verranno rimossi.

In ENVI Zoom, i poligoni che subiscono lo smoothing devono conservare almeno cinque vertici e le polilinee devono avere almeno quattro vertici.

Il programma trasforma il valore dello Smoothing Threshold in map units prima di applicare lo smoothing. Il valore di default dello Smoothing Threshold è pari a 1 pixel. È espresso in valori interi per indicare il livello di generalizzazione, il massimo valore è 99 nonostante valori superiori a 10 siano spesso inusuali. Valori alti riducono i dettagli delle polilinea e dei poligoni. Un chiaro esempio di come sia influente il valore dello Smoothing Threshold è dato dalla figura sottostante.



Cliccando Next gli shapefiles vettoriali e/o le immagini classificate vengono creati. Il Feature Extraction procede con il Reporting step.

### REPORTS E STATISTICHE

Il programma termina il processo di classificazione riportando un sommario dei parametri e delle impostazioni adottate. A titolo di esempio ne viene riportato uno per mostrare tutte le voci che comprende:

```
report - Blocco note
File Name: can_tmr.img
Segment Scale Level: 50.0
Merge Level: 0.0
Refine: No Thresholding

Attributes Computed:
  Spatial
  Spectral
  Texture
  Color Space
    TM Band 3 (0.6600)
    TM Band 2 (0.5600)
    TM Band 1 (0.4850)
  Band Ratio
    TM Band 3 (0.6600)
    TM Band 4 (0.8300)

Classification: Supervised (K Nearest Neighbor)

Export options:
  Vector Output File: C:\Documents and Settings\vector.shp
  Feature Info:
    Feature_1 Type: Polygon
    Feature_2 Type: Polygon
    Feature_3 Type: Polygon
    Feature_4 Type: Polygon
    Feature_5 Type: Polygon
  Smoothing: Threshold of 1

Feature Name   Feature Count   Total Area   Mean Area   Min Area   Max Area
Feature_1     691             112385.00   162.64     1.00      92843.50
Feature_2     2               29513.00   14756.50   11784.00  17729.00
Feature_3     1               45083.00   45083.00   45083.00  45083.00
Feature_4     1               15117.50   15117.50   15117.50  15117.50
Feature_5     455             53712.50   118.05     2.00      9211.50
```

Se si è soddisfatti dei risultati della classificazione si può terminare il processo cliccando il tasto finish oppure si ha la possibilità di percorrere al ritroso il processo e apportare cambiamenti.

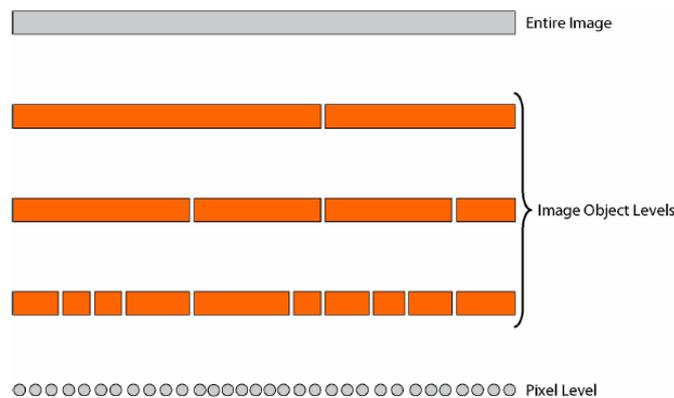
## 4.2 Definiens Developer 7.0

Definiens Developer è un programma che consente l'analisi delle immagini con svariate potenzialità. L'interfaccia grafica permette un rapido sviluppo di regole, moduli e componenti, ad ogni passo si ha la visualizzazione dei risultati così l'analisi dell'immagine avviene interattivamente. Lo studio di tutte le variabili avviene con un linguaggio logico basato sull'interpretazione visiva dell'operatore.

La segmentazione divide l'immagine in aree parziali che presentano differenti caratteristiche dette IMAGE OBJECTS. Gli image objects vengono classificati secondo particolari criteri; Definiens Developer presenta un workflow iterativo: si ha una procedura iterativa di segmentazione (oppure reshaping) e classificazione.

Un'immagine è un set di dati raster ed è formata da almeno un layer; i layer più comuni sono quelli del Rosso del Verde e del Blu. Questo programma rappresenta le immagini con le scenes che consistono in uno o più image layers o canali. Le scene vengono gestite dai PROJECTS che sono contenute a loro volta nel WORKSPACE

Durante l'analisi dell'immagine, la scene viene suddivisa in image objects, ovvero in gruppi di pixels correlati nella stessa scena che definiscono una regione con le informazioni che contengono.



Ogni image object è collegato ai suoi vicini formando una rete e sono organizzati in livelli. Un image object level viene usato come un'area interna di lavoro per analisi dell'immagine.

Durante l'analisi dell'immagine possono essere creati livelli multipli degli image objects che giacciono sopra al livello dei pixels.

Va sottolineata la differenza che si ha tra gli image object levels e image layers.

Gli image layers rappresentano le informazioni già esistenti nell'immagine prima che venga importata all'interno del programma. Al contrario, gli image object levels contengono gli image objects e sono utilizzati come aree di lavoro interno.

Un image object level può essere creato dalla segmentazione di un livello sottostante o di un livello già esistente di image object, oppure da una duplicazione di un esistente image object level.

Sugli image object è possibile eseguire operazioni come: classificazione, reshaping ed estrazione di informazioni.

Questo software col termine feature intende un attributo che contiene informazioni riguardanti gli objects d'interesse. Si hanno due tipologie di feature:

- Image Object features: sono relative agli image objects. Le object features descrivono le proprietà di un image object di tipo spettrale, forma, gerarchia... Si possono estrarre maggiori informazioni da regioni dell'immagine rispetto quelle raccolte da pixels singoli come: colore, forma e tessitura associata alle regioni. Inoltre grazie alla struttura a livelli si hanno informazioni aggiuntive come le relazioni tra objects vicini e il numero di subobjects di una data classe.

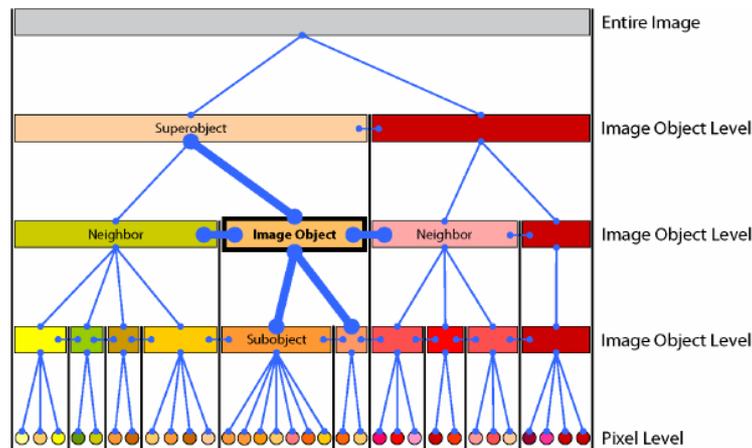
- Global features: considerano gli image objects nel loro complesso. Ad esempio il numero di image objects appartenenti ad una certa classe.

Per classe viene intesa una categoria di image objects che hanno una loro etichetta o nome. Le classi descrivono un significato semantico; vengono raggruppate in una struttura ad albero detta class hierarchy

### Gerarchia degli Image Object

Durante l'analisi dell'immagine possono essere creati image object levels che vengono posizionati sopra il pixel level. L'Image Object Hierarchy funge da magazzino per tutti gli image objects levels che a loro volta contengono i diversi image objects.

Tutti gli image objects sono organizzati in una rete gerarchica che è formata da uno o più image object levels, dalla risoluzione maggiore alla minore. Gli image objects sono collegati tra loro sia verticalmente che orizzontalmente, in modo tale che sono sempre noti i vicini, i superobject che si trovano al livello superiore e i subobjects che sono sull'Image Object Level inferiore, come viene mostrato in figura:



Per garantire l'unicità di queste relazioni, nessun image object può avere più di un superobject ma può avere molteplici subobjects. Il confine di un superobject è formato dal bordo dei suoi subobjects.

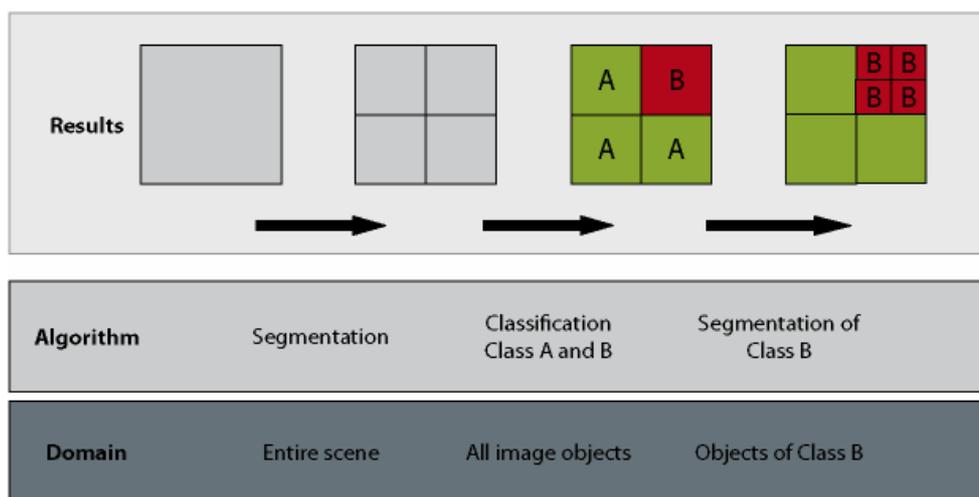
Il processo per la ricerca degli objects in un'immagine consiste in una sequenza di procedure di segmentazione e classificazione; questa sequenza viene definita nel rule set che consiste in una lista di processi rappresentanti dei comandi organizzati in una struttura ad albero.

Per Ruleware viene intesa una porzione di software applicabile ad un rule set, ad un'action, o ad una solution. Una solution è processo specifico di analisi dell'immagine già pronto all'uso e costituito da actions. Un'action è un blocco predefinito dell'image analysis solution. Quando viene configurata consiste in un set di processi con parametri definiti, sono disponibili delle librerie di actions oppure possono essere create per lo specifico scopo. Il Rule set rappresenta il codice del ruleware unito ad impostazioni o funzioni. Un singolo process è l'unità elementare di un rule set creato per l'analisi dell'immagine. Col termine process viene indicato sia il singolo processo che una sequenza di processi. Le principali parti funzionali di un singolo processo sono l'algoritmo e il dominio degli image objects. Un singolo processo consente l'applicazione di uno specifico algoritmo su di una specifica regione d'interesse dell'immagine. I processi possono presentare un numero svariato di processi figli. Sistemare processi contenenti differenti tipi di algoritmi permette all'utente di costruire un blocco per l'analisi sequenziale dell'immagine. L'algoritmo esegue l'operazione che rappresenta il processo, può generare image objects, eseguire il merge o la separazione di image objects, la classificazione... Le due principali funzioni degli algoritmi sono quella di generare o modificare gli image objects e quella di classificazione degli image objects. Per l'analisi dell'immagine saranno necessari e sono disponibili anche queste categorie di algoritmi:

- Segmentation algorithms
- Classification algorithms
- Variables operation algorithms
- Reshaping algorithms
- Level operation algorithms
- Interactive operation algorithms
- Sample operation
- Image layer operation algorithms
- Thematic layer operation algorithms

- Export algorithms
- Workspace automation algorithms
- Process related operation

Il dominio degli Image Objects identifica la region of interest (ROI) sulla quale verrà eseguito l’algoritmo, relativo al processo, tenendo in considerazione l’image object hierarchy. Come mostrato nella figura sottostante, esempi di image object domains sono l’intera immagine, un singolo object level o tutti gli image objects di una specifica classe.



Le immagini possono essere analizzate in singoli progetti o in workspace che gestiscono più di un progetto simultaneamente; è consentito caricare nello stesso progetto immagini e layers tematici che presentano diverse risoluzioni spaziali, per sovrapporle è necessario ricorrere al resampling sull’immagine a minor risoluzione. Per i dati che non sono georeferenziati vengono create automaticamente delle coordinate virtuali in unità pari ad 1pixel.

I dati del progetto posso essere visualizzati a monitor in diversi modi: in scale di grigio, a colori RGB, in falso colore... e sono possibili correzioni sugli istogrammi delle immagini.

Una volta creati gli objects sono disponibili diverse visualizzazioni dei loro bordi detti outlines.

### **Sviluppare un rule set**

Lo scopo dell'analisi con DEFINIENS DEVELOPER è la classificazione degli image objects. L'analisi è basata su questo semplice principio: la ricerca automatica di objects in un'immagine digitale parte col taglio dell'immagine in pezzi chiamati image objects primitivi che devono essere abbastanza omogenei al loro interno e presentare forma e dimensioni idonei all'applicazione di ulteriori processi. Questo step è detto segmentazione. La fase successiva, detta classificazione, consiste nell'etichettare gli image objects in accordo agli attributi, si procede ricorsivamente con un'ulteriore classificazione che la rende più significativa terminando quando gli image objects desiderati sono isolati.

Comunemente la segmentazione è la suddivisione di un'immagine digitale in piccole porzioni in accordo col criterio dato, in Definiens viene indicata con questa parola qualunque operazione che crea image objects quindi anche un'operazione di merge.

L'obiettivo è convertire le informazioni contenute nell'immagine nel processo di analisi della stessa. Col termine feature vengono chiamati gli attributi degli image objects, ci sono due principali tipologie:

**OBJECT FEATURES:** sono attributi dell'image object come l'area

**GLOBAL FEATURES:** non sono relative ad un singolo image object, ma riguardano condizioni generali come il numero di image objects appartenenti ad una certa classe

Un rule set può essere di due tipi:

1 **ROBUSTO:** è composto da processi corti che lavorano veloci e sono creati ad hoc per un'immagine

2 STEP BY STEP: è più flessibile e può essere applicato ad una serie di immagini che presentano anche differenze sostanziali.

Il workflow tipico consiste in tre attività iterative:

uso di più metodologie per ricavare informazioni sugli image objects

editare una o più linee del rule set

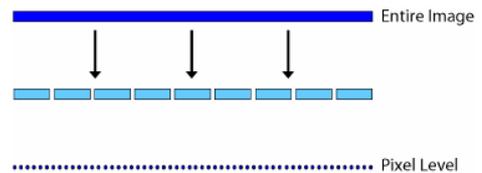
eseguire soltanto la parte del rule set modificata per testarla

Il rule set viene costruito combinando singoli processi in sequenza.

Con la segmentazione si ottengono gli image objects primitivi, che non sono ancora classificati, devono essere delle dimensioni maggiori possibili ma abbastanza piccole per mostrare i contorni della struttura che ci interessa. Sono possibili due strategie di segmentazione:

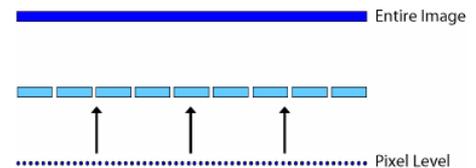
**TOP DOWN SEGMENTATION:**

taglia gli objects creandone dei più piccoli, si può applicare anche all'immagine intera. Esempi sono la segmentazione chessboard e quadtree.



**BOTTOM-UP SEGMENTATION**

Assembla oggetti per crearne altri più grandi, è possibile applicarla al livello dei pixels.



Sono possibili due metodologie di classificazione degli image objects che possono anche essere combinate per rinforzare i risultati:

1) **PROCESS-BASED CLASSIFICATION:** consente di combinare svariati step di classificazione degli image objects in un processo. Gli algoritmi di classificazione analizzano gli image objects e assegnano loro una classe in accordo ai criteri definiti, sono:

**ASSIGN CLASS:** assegna le classi in accordo alle feature;

**CLASSIFICATION:** l'assegnazione segue la descrizione delle classi;

**HIERARCHICAL CLASSIFICATION:** utilizza la struttura gerarchica delle classi;

**ADVANCED CLASSIFICATION ALGORITHMS:** sono algoritmi progettati per creare uno specifico comando di classificazione.

2) **CLASS DESCRIPTION:** consente l'uso del Fuzzy Logic; è necessario stabilire una gerarchia delle classi con le seguenti tipologie di espressioni e metodi:

**SOGLIE:** utile quando le classi possono essere differenziate chiaramente da una feature

**MEMBERSHIP FUNCTIONS:** le membership functions offrono una relazione tra i valori della feature e il grado di appartenenza ad una classe. Vengono definite dai valori assunti dai bordi.

Il Fuzzy Logic è un approccio matematico che quantifica un'incertezza in grado di emulare il pensiero umano. Questo metodo considera possibile per ogni object più di una classe, verrà assegnata quella che presenta il maggior valore dalla membership function.

Button	Function Form	Button	Function Form
	Larger than		Linear range (triangle)
	Smaller than		Linear range (triangle inverted)
	Larger than (Boolean, crisp)		Singleton (exactly one value)
	Smaller than (Boolean, crisp)		Approximate Gaussian
	Larger than (linear)		About range
	Smaller than (linear)		Full range

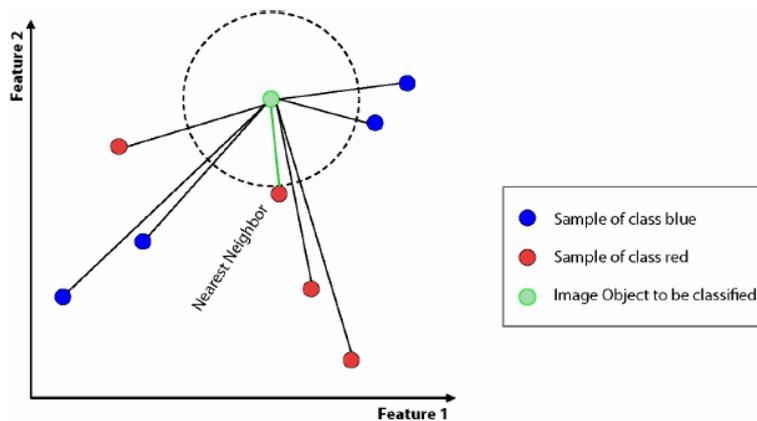
**SIMILARITIES:** fa uso di similitudini nei confronti delle classi

**ESEMPI PER LA CLASSIFICAZIONE NEAREST NEIGHBOR:** richiede esempi tipici rappresentativi di ogni classe.

La procedura consiste in due step principali:

istruire il sistema fornendogli alcuni image objects come esempi,

classificare gli image objects tramite l'uso dei nearest sample neighbors.



### Controllo dei risultati

Al termine del processamento dell'immagine, è buona norma controllare i risultati e valutare se il rule set è ottimale. Il programma consente diverse possibilità, compreso l'editing manuale per far correggere direttamente all'utente i risultati errati.

I risultati dell'analisi dell'immagine possono essere consultati sia in forma numerica di tipo statistico che visiva con la classificazione degli image objects.

Per la verifica dell'accuratezza di una classificazione si può usare il metodo della matrice di confusione, nelle righe si ha l'accuratezza del classificatore dal punto di vista di chi usa la mappa classificata, lungo le colonne si ha l'accuratezza della classificazione dal punto di vista di chi produce la classificazione.

User Class \ Sample	Impervious General	Grassland General	Woodland General	Waterbodies	Sum
<b>Confusion Matrix</b>					
Impervious General	7	0	0	0	7
Grassland General	0	0	0	0	0
Woodland General	0	0	0	0	0
Waterbodies	0	0	0	8	8
unclassified	0	12	15	0	27
Sum	7	12	15	8	
<b>Accuracy</b>					
Producer	1	0	0	1	
User	1	undefined	undefined	1	
Hellden	1	0	0	1	
Short	1	0	0	1	
KIA Per Class	1	0	0	1	
<b>Totals</b>					
<b>Overall Accuracy</b>	<b>0.3571</b>				
<b>KIA</b>	<b>0.3131</b>				

Overall Accuracy rappresenta l'accuratezza complessiva; e data dal numero di siti correttamente classificati diviso per il numero totale di punti. KIA è una statistica sulla diagonale e sui termini marginali della matrice di confusione.

### **Esportazione dei risultati**

I risultati dell'analisi possono essere esportati in formato vettoriale o raster ed eventualmente correlati dalle relative statistiche.

## **4.3 Considerazioni**

I due programmi hanno le stessa finalità. Dopo questa breve descrizione è evidente come Envi Feature Extraction sia di più immediato utilizzo e comprensione; proprio per il suo schema rigido consente all'utente poche personalizzazioni e risultati non sempre accettabili. Il software Definiens Developer al contrario non è di uso intuitivo, richiede una buona conoscenza della sua architettura, consente personalizzazioni minuziose. La libertà concessa all'utente durante l'analisi dell'immagine permette agevolmente di ottenere buoni risultati, che sono comunque rimaneggiabili al termine del processamento, anche manualmente, a differenza di Envi.

Personalmente ritengo Envi Feature Extraction idoneo ad una prima analisi veloce; mentre Definiens Developer è più indicato ad analizzare immagini che presentano una maggior complessità nella comprensione, che richiedono risultati dettagliati con molte classi e sottoclassi, necessariamente è richiesto un tempo di lavoro superiore all'utente.

## Capitolo V

# Sperimentazione su casi di studio

Successivamente allo studio degli aspetti teorici, è stata svolta una sperimentazione su alcuni casi applicativi per testare le diverse metodologie di classificazione. In questo capitolo vengono riportati i risultati ottenuti dall'analisi di immagini utilizzando i software descritti precedentemente (Envi 4.5, Definiens Developer 7.0). Gli ambiti di studio prevedono due tipologie distinte di dato: nel primo caso verranno studiati oggetti di piccole dimensioni, nel secondo verranno analizzate aree urbane con la finalità di estrarne gli edifici.

Nel primo ambito di studio l'obiettivo è individuare una procedura automatica che consenta un conteggio meccanico su di un'immagine a colori RGB non satellitare di frammenti di ceramiche su un'area archeologica e che sia capace di distinguere le diverse tipologie presenti. Il successo di questo metodo costituirebbe una soluzione alternativa al conteggio manuale in sito tuttora adottato.

Nel secondo ambito (costituito da dataset diversi relativi a differenti aree) i dati a disposizione sono immagini satellitari ad alta risoluzione acquisite dal sensore QuickBird su aree urbane. La classificazione object-based dovrebbe fornire buoni risultati per questa tipologia di dati caratterizzati da una bassa risoluzione spettrale e da una elevata risoluzione spaziale che mette in evidenza l'eterogeneità degli elementi che lo contraddistinguono; in questo caso oggetto di interesse sono gli edifici.

Nell'ambito della classificazione object-based, per ottenere risultati soddisfacenti, come prima cosa è buona norma analizzare visivamente l'immagine per cogliere gli elementi ritenuti idonei a rappresentare gli oggetti di studio. Il rule set deve essere la trasformazione in un linguaggio di calcolo del processo d'interpretazione dell'immagine compiuto dalla mente umana. Nella classificazione supervisionata l'analisi può essere schematizzata in questi punti:

- 1 visione dell'immagine;
- 2 scelta dei dati idonei;
- 3 sviluppo della strategia;
- 4 trasportare la strategia nel rule-set;
- 5 controllo dei risultati;
- 6 perfezionamento della strategia e del rule-set quando necessario;
- 7 esportazione dei risultati.

Queste fasi sono state seguite per entrambi gli ambiti di studio.

### **5.1 Conteggio di reperti archeologici di piccole dimensioni**

La sperimentazione qui descritta riguarda la missione archeologica congiunta italo-turcomanna nel conoide alluvionale del fiume Murghab (Turkmenistan) (Cerasetti, 2009). Il progetto archeologico nella Repubblica del Turkmenistan è basato sul protocollo di Cooperazione Scientifica stipulato tra l'Istituto Italiano per l'Africa e l'Oriente (IsIAO) di Roma ed il Ministry of Culture and Tv and Radio Broadcasting of Turkmenistan, con la collaborazione del Dipartimento di Archeologia dell'Università di Bologna. La direzione è del Dr. Mukhammed A. Mamedov da parte turcomanna e della Dr.ssa Barbara Cerasetti da parte italiana. Numerosi sono i partner nazionali ed internazionali. Scopo primario del progetto è lo studio del fenomeno urbano nel Bronzo e nel Ferro attraverso lo scavo archeologico e l'analisi territoriale.



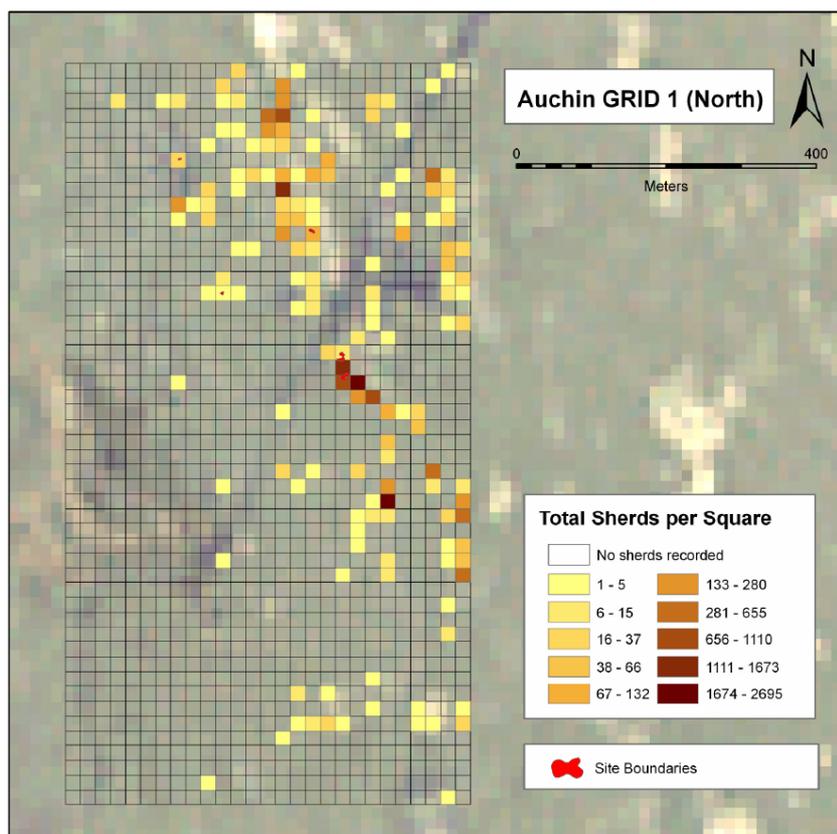
Immagine satellitare del Turkmenistan con evidenziata l'area del conoide alluvionale del fiume Murghab

Tra le attività della missione in questi anni vi è stata la ricerca, il conteggio e la catalogazione di frammenti di reperti ceramici di superficie presenti per estensioni molto ampie del territorio oggetto di studio.



Dispersione di frammenti ceramici in superficie: sito n. 1529 visto da Sud

La sperimentazione qui descritta risponde all'esigenza della missione archeologica di mettere a punto una metodologia per la mappatura, in numero e tipologia di materiale, di tali frammenti, in alternativa alla tecnica fin qui seguita, basata sul conteggio e la raccolta manuale all'interno di celle di un grigliato regolare sul terreno.



Griglia Nord nell'area Auchin.

### **Dati a disposizione**

Relativamente a questo caso di studio sono a disposizione due immagini a colori di dimensioni 1200x1600 pixel, gentilmente messe a disposizione dalla Dr.ssa Barbara Cerasetti. E' stato inserito un elemento metrico nelle immagini per un dimensionamento di massima del contenuto; è evidente che la sola riproduzione della fettuccia metrica non può tuttavia essere sufficiente ad ottenere

risultati di una certa accuratezza, nemmeno con un semplice radrizzamento che avrebbe richiesto la presenza di almeno quattro punti noti.

L'immagine 1201 presenta una distribuzione spaziale dei frammenti ben distribuita su tutta l'immagine, la dimensione media è di circa due centimetri. L'immagine 1219 invece mostra una porzione di spazio più ampia dove compaiono pochi frammenti di dimensioni più importanti rispetto ai precedenti, di ordine anche decimetrico.

Una prima classificazione non supervisionata può essere utile per mettere in risalto il comportamento spettrale delle varie classi; è però una tecnica attendibile soprattutto per immagini multispettrali e iperspettrali e non risulta pertanto del tutto significativa per immagini con le sole bande RGB.



*Immagine 1201*



*Immagine 1219*

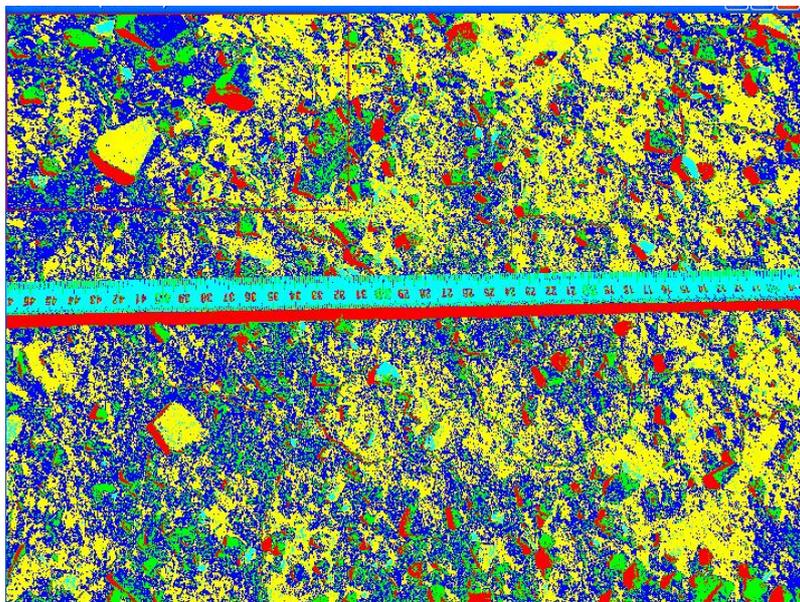
### **5.1.1 Estrazione frammenti dall'immagine 1201**

Sono state definite a titolo sperimentale quattro classi di interesse in base ai colori predominanti: frammenti chiari, frammenti rossi chiari, frammenti rossi, frammenti scuri. Va notato che gli oggetti d'interesse sono poco distinguibili spettralmente dal resto dell'immagine.

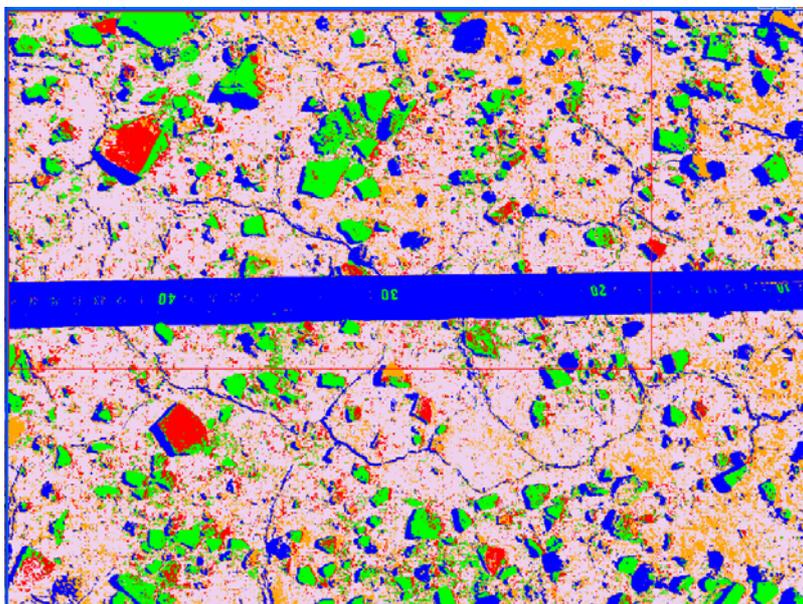
La prima classificazione effettuata è stata quella di tipo non supervisionato, tramite il software Envi, al fine di avere una prima caratterizzazione spettrale. Essa ha confermato come vaste porzioni dell'immagine siano poco differenziabili dagli oggetti d'interesse e vengano attribuite a frammenti della categoria rossi chiari. È stata eseguita con l'algoritmo *k-means* impostando 5 classi e 10 iterazioni. Il risultato di questa classificazione non è soddisfacente.

Risultati migliori possono essere ottenuti con una classificazione supervisionata; l'algoritmo di assegnazione a massima verosimiglianza produce infatti una netta distinzione dei frammenti d'interesse dal resto dell'immagine. Per

potere applicare la classificazione supervisionata è stato necessario definire le ROI, ovvero sono state individuate alcune porzioni d'immagini alle quali è stato associato un significato semantico che viene utilizzato durante la fase di classificazione.



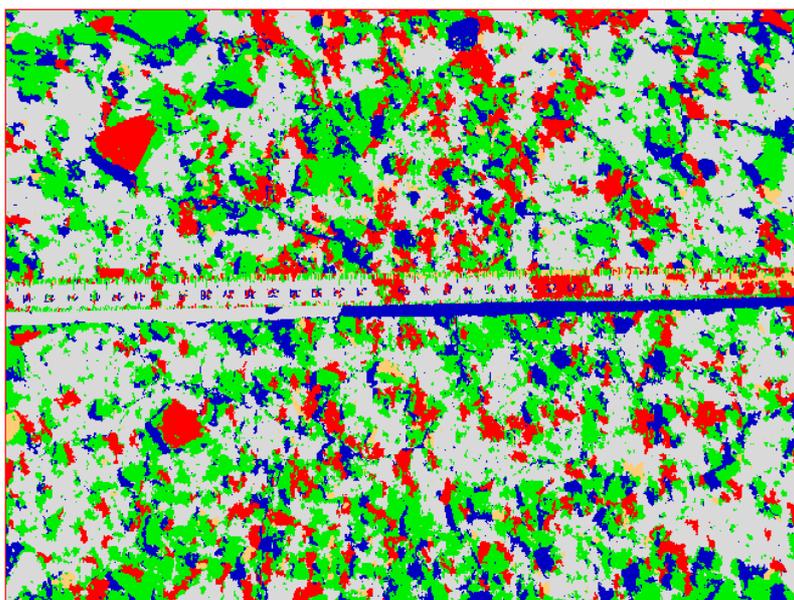
Classificazione non supervisionata pixel-based con algoritmo k-means



Classificazione supervisionata pixel-based con algoritmo di massima verosimiglianza

L'immagine è stata classificata anche con la metodologia object-based. I risultati ottenuti col modulo Feature Extraction di Envi non sono all'altezza delle

aspettative: la qualità di questa classificazione è addirittura inferiore alla classificazione pixel-based con algoritmo di massima verosimiglianza. Il processo è stato eseguito impostando: scala segmentazione pari a 30, merge level 80.0, nessun thresholding, classificazione tramite samples con algoritmo di classificazione supervisionata K-Nearest Neighbor.



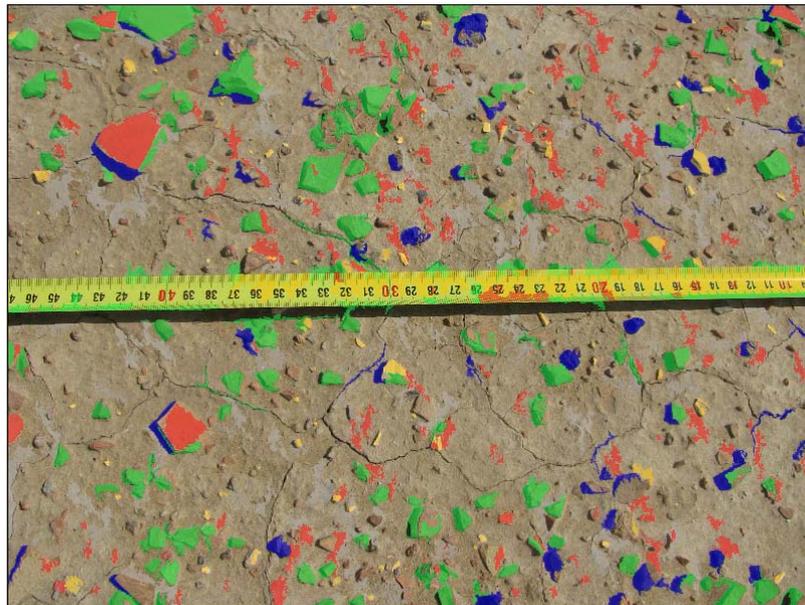
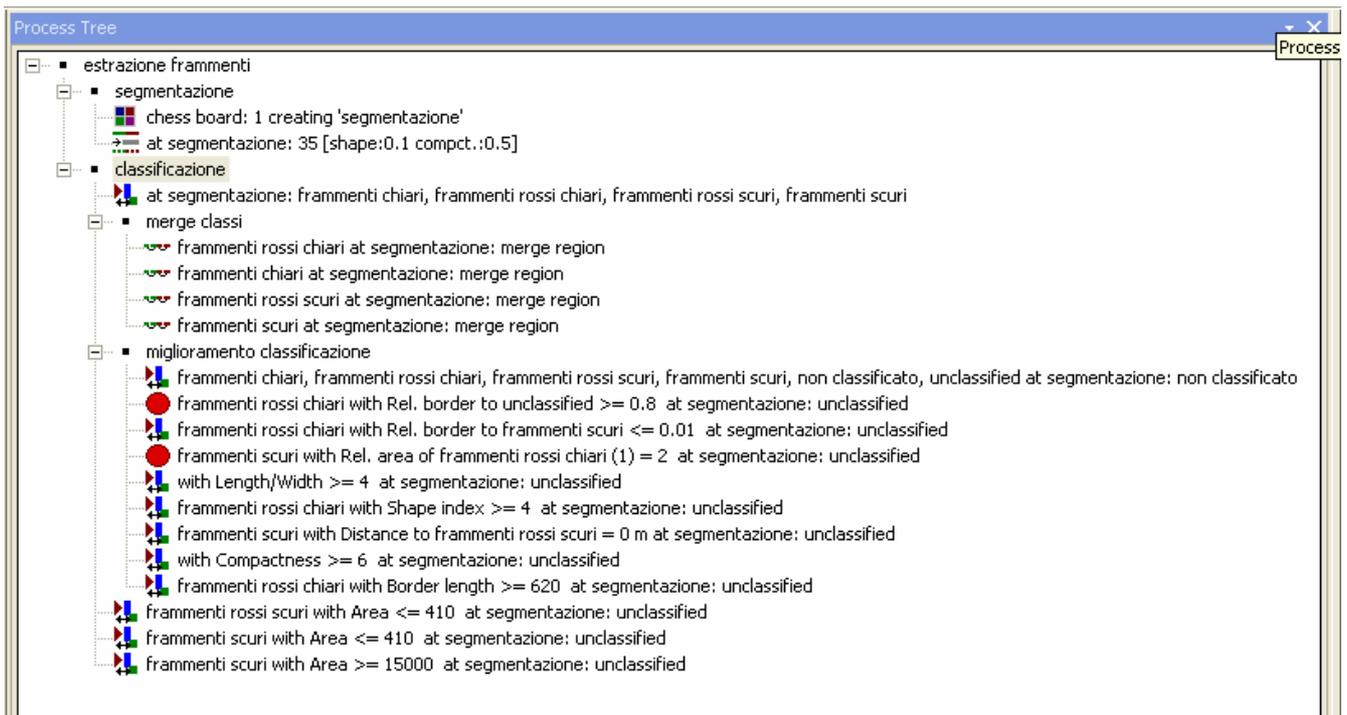
Classificazione supervisionata eseguita con Envi Feature Extraction

Con il software Definiens è stato elaborato un processo che presenta due fasi ben distinte: una iniziale per la segmentazione e una più elaborata per la classificazione dell'immagine. È stata effettuata una segmentazione del tipo multirisoluzione caratterizzata da un fattore di scala 35, schape 0,1 e compactness 0,5; questi parametri stabiliscono le dimensioni e la forma degli image objects. La classificazione utilizzata è del tipo supervisionato con descrizione delle classi, seguita da una fase di merge per accorpare gli objects appartenenti alla stessa classe. Viene creata anche la classe *non classificato* per identificare il terreno ed isolarlo dagli oggetti di interesse. A seguito della segmentazione sono stati ricercati gli attributi che definiscono al meglio i frammenti da conteggiare, i loro valori vengono usati per la descrizione delle classi.

Le classi di studio sono le stesse quattro già introdotte; nella seguente tabella vengono riportate gli attributi che le contraddistinguono:

classe	colore	descrizione
<i>Frammenti chiari</i>	giallo	Brightness $\geq 164$
<i>Frammenti rossi chiari</i>	rosso	$151 \leq \text{Brightness} \leq 185$
<i>Frammenti rossi scuri</i>	verde	$0.5 \leq \text{Max. Diff.} \leq 2.8$
<i>Frammenti scuri</i>	blu.	Brightness $\leq 100$
<i>Non classificato</i>	grigio	$3.8 \leq \text{Shape Index} \leq 4.1$

La classificazione è stata eseguita con l’algoritmo *classification* che opera sull’intera immagine utilizzando la descrizione delle classi. Essendo il fine ultimo il conteggio automatico dei frammenti, è previsto un merge-objects che fonde tutti gli objects adiacenti che appartengono alla stessa classe. Al primo controllo i risultati della classificazione interpretano abbastanza bene i frammenti rossi chiari, rossi scuri e chiari, ma sono presenti errori di misclassificazione per la classe dei frammenti scuri che include le ombre e le fratture nel terreno. A seguito della prima procedura di classificazione, è necessario tentare un miglioramento dei risultati sfruttando le features contestuali; in particolare sono state fatte considerazioni sull’area e relazioni tra i bordi degli objects contigui. La figura sottostante illustra il processo sviluppato per l’estrazione dei frammenti dall’immagine 1201.



Classificazione supervisionata object-based eseguita con Definiens Developer

Quando i risultati della classificazione non sono più migliorabili si passa alla fase del conteggio dei frammenti. Il software mette a disposizione dell'utente una grande varietà di features, in particolare in questo caso è interessante ricercarne una che indichi il numero di oggetti classificati per ogni classe. Questo requisito è soddisfatto dalla *Scene feature/Class Related: number of classified objects*.

Nella seguente figura sono riportati i risultati del conteggio automatico basato sui risultati della classificazione.

Image Object Information	
Feature	Value
<b>Class-Related</b>	<b>Number of classified objects</b>
frammenti chiari	109
frammenti rossi chiari	154
frammenti rossi scuri	124
frammenti scuri	51

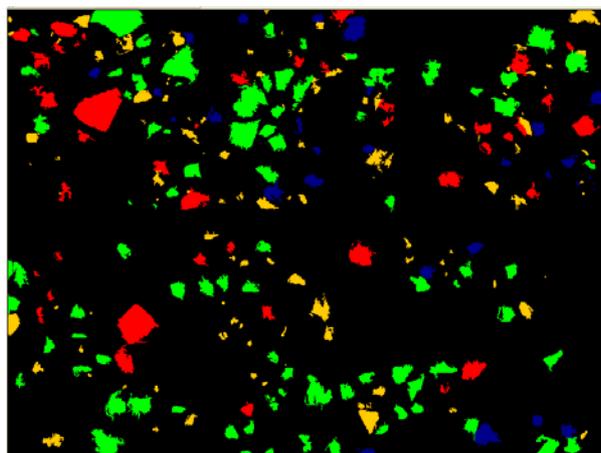
Per saggiare la validità dei risultati ottenuti in automatico è stato fatto un conteggio manuale che viene assunto come verità a terra; la tabella sottostante riporta una stima dell'accuratezza in termini di objects:

	Frammenti			
	Chiari	Rossi chiari	Rossi scuri	Scuri
Verità a terra	128	170	126	31
Classificazione	109	154	124	51
Accuratezza	0.851562	0.905882	0.984127	0.259843

Il conteggio automatico risulta attendibile per i frammenti rossi scuri (98%) e per i frammenti rossi chiari (90%), è abbastanza buono per i frammenti chiari (85%) e per nulla attendibile per i frammenti scuri (25%).

È stata effettuata un'analisi dell'accuratezza della classificazione in termini di pixels utilizzando il software Definiens Developer 7. L'accuratezza della classificazione viene valutata tramite la matrice d'errore che necessita della conoscenza della reale natura dei pixels. Per dare questa conoscenza è necessario creare un secondo progetto caratterizzato dagli stessi parametri di segmentazione,

classificare manualmente gli objects e salvare quest'informazione in un'immagine che viene detta TTA Mask.



TTA Mask

Per procedere alla valutazione dell'accuratezza vengono correlate le classi tra la TTA Mask e l'immagine classificata.

L'immagine è formata da 1'920'000 pixels, la seguente matrice d'errore in termini di pixels ne mostra quantitativamente e qualitativamente la ripartizione tra le classi.

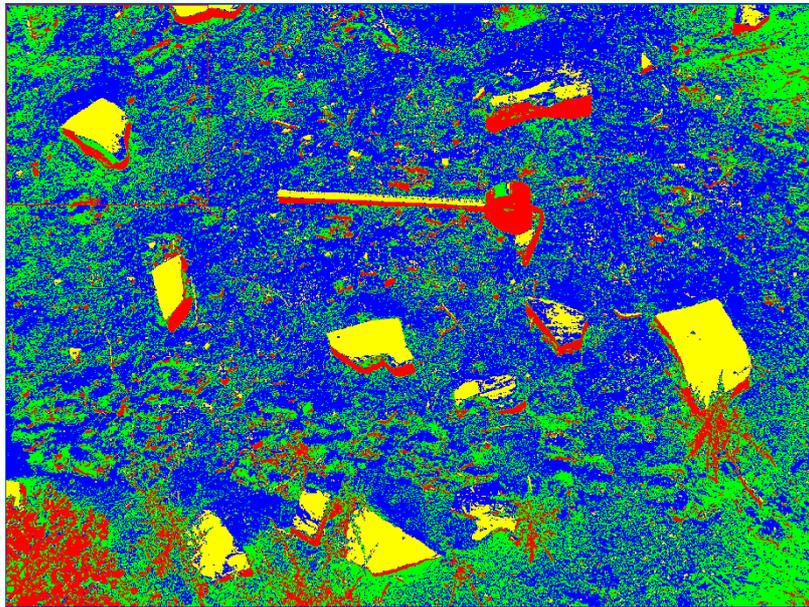
User \ Referenc...	chiari	rossi chiari	rossi	scuri	Sum
<b>Confusion Matrix</b>					
chiari	9958	427	60	40	10485
rossi chiari	2131	14854	0	0	16985
rossi	1891	4371	90827	275	97364
scuri	1067	1257	1900	13614	17838
unclassified	20480	14612	28309	8558	71959
Sum	35527	35521	121096	22487	
<b>Accuracy</b>					
Producer	0.2802939	0.4181752	0.75	0.6054165	
User	0.9497377	0.8745364	0.9328602	0.7632022	
Hellden	0.4328436	0.5658020	0.8315206	0.6752139	
Short	0.2761968	0.3945076	0.7116263	0.5096777	
KIA Per Class	0.2433295	0.3681752	0.5425065	0.5696500	
<b>Totals</b>					
<b>Overall Accuracy</b>	<b>0.6022103</b>				
<b>KIA</b>	<b>0.443</b>				

Si ricorda che sulla diagonale si ha il numero di pixels correttamente classificati mentre gli elementi fuori dalla diagonale sono attribuiti ad una classe sbagliata. L'accuratezza totale pari a 0,6022103 è un valore non del tutto soddisfacente perché soltanto il 60% dei pixels sono correttamente classificati.

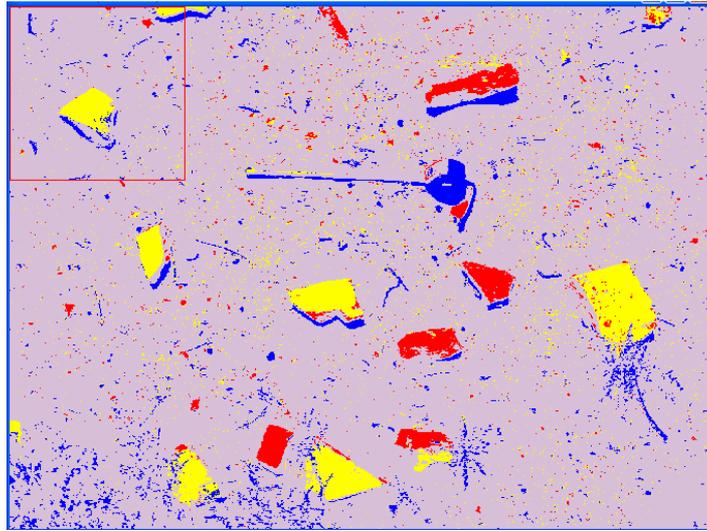
Purtroppo il risultato può essere attribuito alla scarsa distinzione tra i frammenti rossi chiari e il terreno.

### **5.1.2 Estrazione frammenti dall'immagine 1219**

Le classi di interesse sono tre: frammenti chiari, frammenti rossi, frammenti scuri; in questa immagine si ha una separazione spettrale maggiore tra le classi. La prima classificazione effettuata è quella di tipo non supervisionata con l'algoritmo *k-means* impostando 4 classi e 10 iterazioni; il risultato conferma la distinzione più marcata tra le classi e i frammenti vengono discretamente identificati. Risultati migliori si attendono con la classificazione supervisionata. L'algoritmo di massima verosimiglianza individua meglio le classi. Per potere applicare la classificazione supervisionata è stato necessario definire le ROI, ovvero sono state individuate alcune porzioni d'immagini alla quale è stato associato un significato semantico che viene utilizzato durante la fase di classificazione.

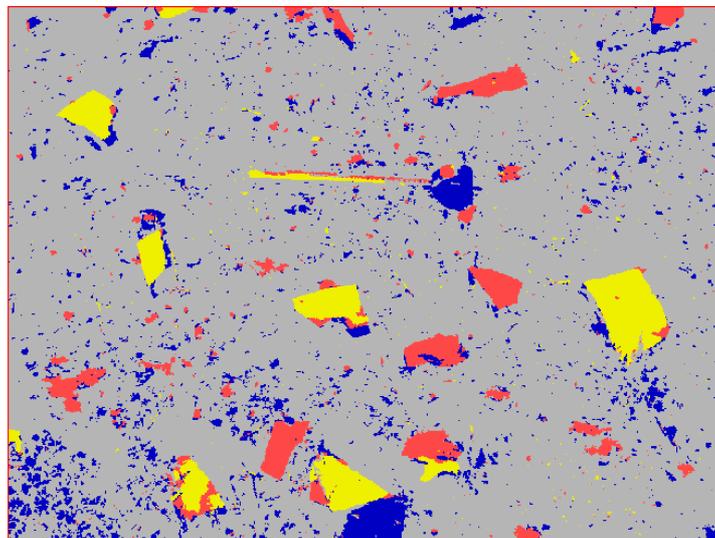


Classificazione non supervisionata pixel-based con algoritmo k-means



Classificazione supervisionata pixel-based con algoritmo di massima verosimiglianza

L'immagine è stata classificata anche con la metodologia object-based. Con questa immagine, i risultati ottenuti col modulo Feature Extraction di Envi rispecchiano maggiormente le aspettative, la qualità di questa classificazione è però paragonabile a quella pixel-based con algoritmo di massima verosimiglianza; non si ha un miglioramento, ed anzi alcuni objects della classe terreno vengono malclassificati in frammenti rossi. Il processo è stato eseguito impostando: scala segmentazione pari a 30, merge level 90.0, nessun thresholding, classificazione tramite samples con algoritmo di classificazione supervisionata K-Nearest Neighbor.



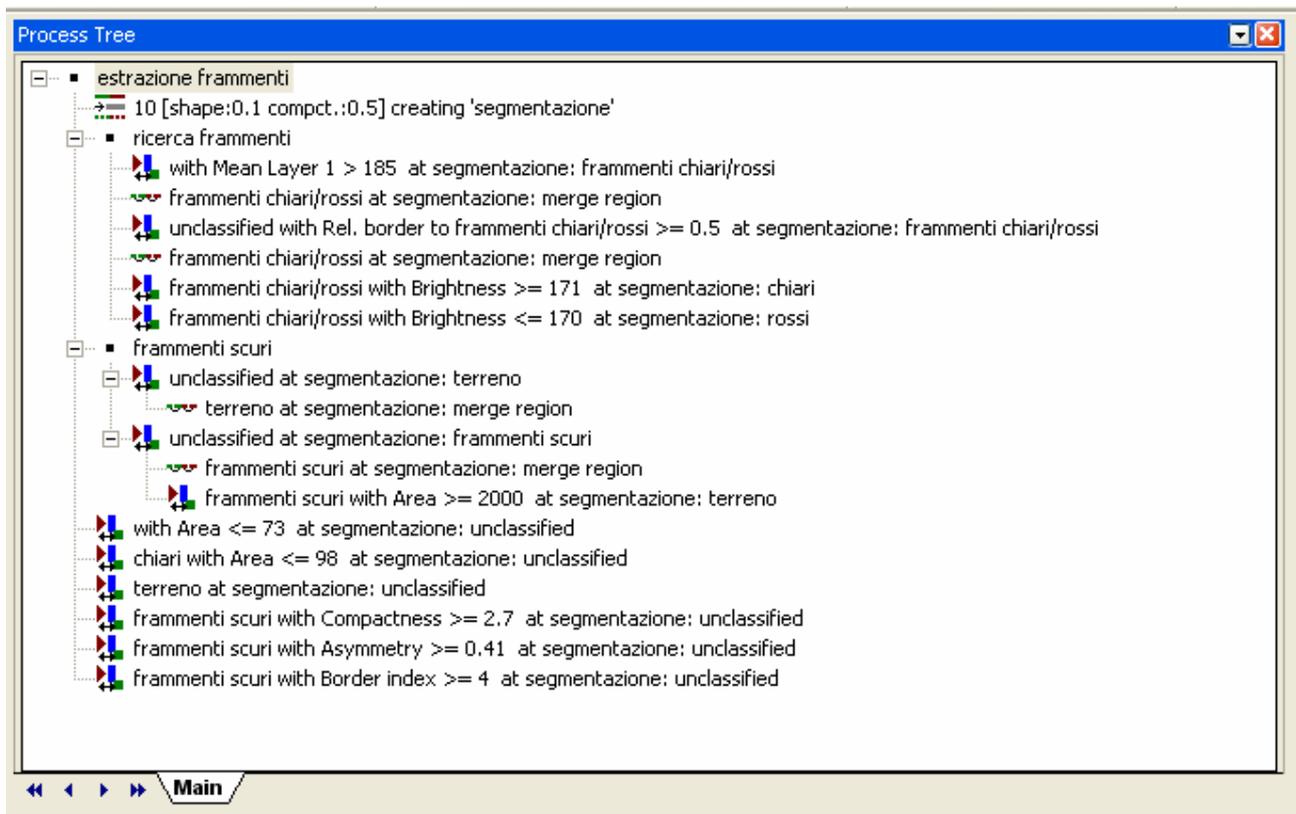
Classificazione supervisionata eseguita con Envi Feature Extraction

Con il software Definiens è stato anche in questo caso elaborato un processo che presenta una fase iniziale per la segmentazione e una più elaborata per la classificazione dell'immagine. È stata effettuata una segmentazione del tipo multirisoluzione caratterizzata da un fattore di scala 10, schape 0,1 e compactness 0,5. La classificazione è stata eseguita con l'algoritmo *assign class* che opera sull'intera immagine utilizzando definizioni di soglie per la descrizione delle classi. Anche in questo caso viene creata la classe *terreno* per isolare lo sfondo dagli oggetti di interesse. È identificato con dei valori di soglia sulla luminosità.

In questa immagine è stato molto semplice individuare i frammenti rossi e chiari: inizialmente sono stati individuati imponendo una condizione di soglia sulla banda del rosso attribuendoli ad una classe fittizia chiamata *rossi/chiari*; la distinzione tra frammenti rossi e chiari è avvenuta all'interno di questa classe fittizia imponendo delle condizioni di soglia sui valori spettrali.

La scelta degli attributi idonei a caratterizzare la classe dei frammenti scuri è stata difficoltosa perché non è stato possibile identificare alcune features capaci di distinguerla marcatamente. A seguito di queste considerazioni si è preferito identificarla selezionando sull'immagine alcuni *samples*, sub set spaziali ai quali è stata attribuita la classe di appartenenza. La classe dei frammenti scuri è stata classificata con un algoritmo Nearest Neighbor.

Al primo controllo i risultati della classificazione interpretano bene i frammenti rossi e chiari, ma sono presenti errori per la classe dei frammenti scuri che include le ombre e le irregolarità del terreno. A seguito della prima procedura di classificazione, è necessario tentare un miglioramento dei risultati sfruttando le features contestuali; in particolare sono state fatte considerazioni sull'area, asimmetria, compattezza e relazioni tra i bordi degli objects contigui. La figura sottostante illustra il processo sviluppato per l'estrazione dei frammenti dall'immagine 1219.



Classificazione supervisionata object-based eseguita con Definiens Developer

Ottenuta la classificazione dell'immagine si hanno a disposizione i risultati del conteggio automatico che sono riportati nella figura sottostante:

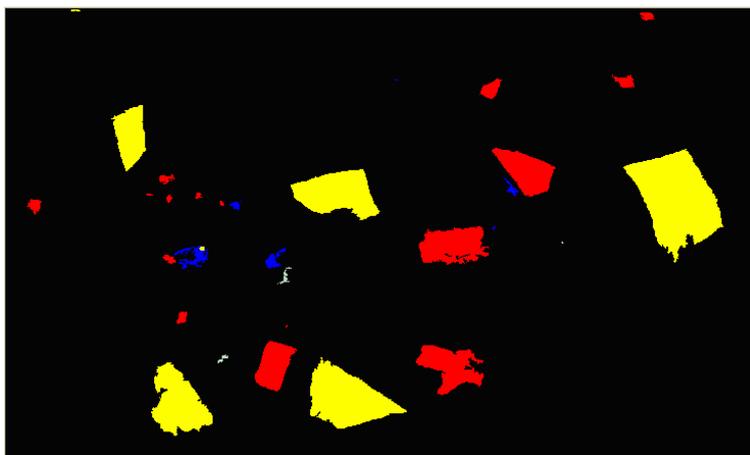
Image Object Information	
Feature	Value
<b>Class-Related</b>	<b>Number of classified objects</b>
chiaro	13
frammenti scuri	127
rossi	34

Anche in questo caso per saggiare la validità dei risultati ottenuti in automatico è stato fatto un conteggio manuale che viene assunto come verità a terra; la tabella sottostante riporta una stima dell'accuratezza in termini di objects:

	Frammenti		
	Chiari	Rossi	Scuri
Verità a terra	12	26	33
Classificazione	13	34	127
Accuratezza	0.923077	0.7647059	0.259843

Il calcolo automatico fornisce un conteggio attendibile per i frammenti chiari (92%), abbastanza buono per i frammenti rossi (76%) e per nulla attendibile per i frammenti scuri (25%).

Anche in questo caso l'accuratezza della classificazione viene valutata in Definiens tramite la matrice d'errore. Bisogna quindi anche per questa immagine classificare manualmente gli objects in un secondo progetto caratterizzato dagli stessi parametri di segmentazione. L'immagine classificata viene confrontata con l'immagine TTA Mask che esprime la verità a terra.



TTA Mask

Anche l'immagine 1219 ha una risoluzione di 1'920'000 pixels. La matrice d'errore mostra quantitativamente e qualitativamente la ripartizione dei pixels tra le classi.

User \ Referenc...	scuri	rossi	chiar	Sum
<b>Confusion Matrix</b>				
scuri	1436	0	0	1436
rossi	0	35259	146	35405
chiar	0	0	71213	71213
unclassified	1714	6175	3583	11472
Sum	3150	41434	74942	
<b>Accuracy</b>				
Producer	0.4558730	0.851	0.9502415	
User	1	0.9958763	1	
Hellden	0.6262538	0.9177371	0.9744860	
Short	0.4558730	0.848	0.9502415	
KIA Per Class	0.4492563	0.7882429	0.8768979	
<b>Totals</b>				
<b>Overall Accuracy</b>	<b>0.9027994</b>			
<b>KIA</b>	<b>0.8143047</b>			

L'accuratezza totale pari a 0,9027994 è un buon risultato nonostante l'evidente errore che si ha anche in questo caso per la classe dei frammenti scuri.

### 5. 1.3 Considerazioni sui risultati

Per l'immagine 1201 il risultato di questa sperimentazione è modesto, più incoraggianti sono i risultati ottenuti dall'immagine 1219. Le dimensioni più grandi dei frammenti, il minor numero di elementi da individuare, la maggior separazione spettrale tra le classi d'interesse e il terreno, la ricerca di sole tre classi e non quattro, sono fattori che hanno reso i risultati di classificazione di questa immagine di qualità superiore all'immagine 1201.

Per entrambe le immagini il conteggio dei frammenti scuri non ha fornito risultati soddisfacenti, si è evidenziato una difficoltà da parte del programma di estrapolare i frammenti di dimensioni minori che sono difficilmente distinguibili dalle ombre e dalle irregolarità del terreno.

Per questi motivi quindi potrebbe essere giustificabile una seconda analisi in cui viene esclusa questa classe.

In generale la procedura di analisi di immagine può comunque considerarsi di notevole interesse, e potrebbe essere approfondita anche in merito allo sviluppo di una piattaforma di acquisizione che consenta di ottenere una copertura sistematica del terreno con strumentazione motorizzata. In una analisi costi-benefici, considerando anche i lunghi tempi e la necessità di personale per effettuare il riscontro diretto sul campo, l'approccio basato su immagini è da considerarsi sicuramente una valida alternativa, pur non raggiungendo la classificazione automatica le stesse performance di successo di una interpretazione visiva.

## 5.2 Estrazioni di edifici da immagini satellitari ad alta risoluzione

Il secondo ambito applicativo si riferisce alla estrazione di edifici ed è dunque a scala territoriale, idoneo per esempio per applicazioni in ambito urbano. In questo caso i dati a disposizione sono immagini satellitari ad alta risoluzione acquisite dal sensore QuickBird e interessano aree urbane. Le immagini a disposizione sono ortorettificate e georeferenziate, quindi presentano le stesse proprietà geometriche di una cartografia.

Il satellite QuickBird è uno dei satelliti commerciali con la più alta risoluzione spaziale disponibile, è caratterizzato da una risoluzione nominale di 0.61m per il sensore pancromatico (al nadir), mentre per la componente multispettrale (quattro bande) la risoluzione nominale è di 2.44m a terra.

Di rilievo sono le caratteristiche che presentano le immagini *pansharpened*. Il *pansharpening* è un'operazione tra immagini che consente la fusione di un'immagine ad alta risoluzione ottenuta in una banda pancromatica con un'immagine multispettrale a bassa risoluzione per ottenere un'unica immagine multispettrale ad alta risoluzione, che abbia i vantaggi in termini di risoluzione geometrica del pancromatico e l'espressività di una mappa composita a colori.

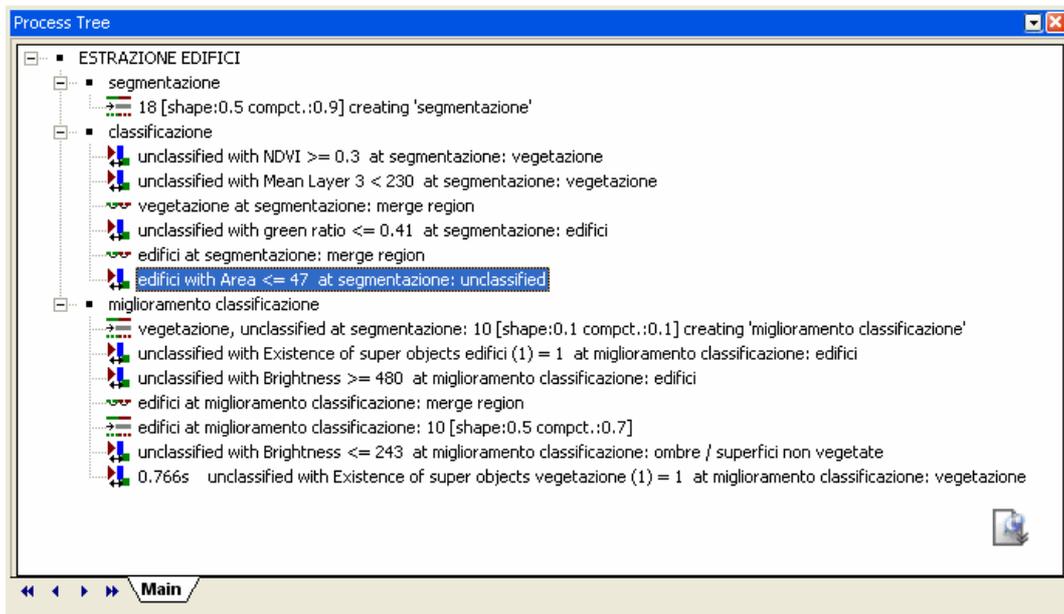
### 5.2.1. Primo caso di studio: Comune di Bologna

L'area di studio riguarda il centro città di Bologna, in particolare il foglio 221 IGM. L'immagine pancromatica è stata rilevata nel luglio 2003, il dato multispettrale è stato acquisito in un momento temporale diverso (2002-2004). Le due immagini hanno subito il *pansharpening* per ottenere i vantaggi dell'elevata risoluzione spaziale del pancromatico e compensarne le carenze d'informazione spettrale.

L'obiettivo è individuare gli edifici tramite una segmentazione che ne delinei il perimetro. Come di consueto il processo ha inizio con una segmentazione. La segmentazione adottata è del tipo multirisoluzione ed è contraddistinta da un parametro scala pari a 18, shape 0.5 e compactness 0.9. Il valore elevato

assegnato alla compattezza è stato deciso per prediligere la forma piuttosto che l'informazione spettrale, perché si vogliono ottenere elementi rettangolari che non presentino irregolarità spiccate. La classificazione è stata eseguita tramite ripetute applicazioni dell'algoritmo *assign class*. Una prima caratteristica in grado di distinguere gli objects all'interno dell'immagine è senz'altro l'indice di vegetazione NDVI; con un valore di soglia pari a 0.3 è possibile distinguere le aree vegetate dal resto dell'immagine. La classe edifici viene individuata ponendo un valore di soglia della banda del rosso  $> 230$ .

Si è riscontrato che un attributo utile a distinguere gli edifici dalla vegetazione oltre all'NDVI è il *green ratio*: un rapporto tra la banda del verde e la somma delle tre bande del visibile; gli edifici sono ben riconoscibili sotto una soglia di questo attributo pari a 0.41. A questo punto buona parte degli edifici risultano classificati, vengono invece rimossi dalla classe gli objects che presentano un'area  $< 47$ .



Albero dei processi

Per tentare di migliorare i bordi degli objects classificati come edifici, si è pensato di eseguire una segmentazione multirisoluzione sulle classi vegetazione e

non classificato, caratterizzata da un parametro di scala 10, shape 0.1, compactness 0.1; ora non si vuole prediligere la forma degli objects, ma si vogliono cogliere le informazioni spettrali. Con questa segmentazione si va a creare un livello sottostante a quello precedentemente classificato; avendo a disposizione 2 livelli aumentano le features a disposizione, non si hanno soltanto le relazioni tra objects vicini ma anche tra sub e super-objects. Il tentativo di miglioramento della classificazione non ha apportato risultati significativi, ma sono state messe in evidenza l'aree di verde della zona. La conoscenza della copertura della vegetazione può essere utilizzata per ricavare indicatori di qualità dell'ambiente urbano.



*Dettaglio dell'immagine classificata*

### 5.2.2 Secondo caso di studio: area di Caselle

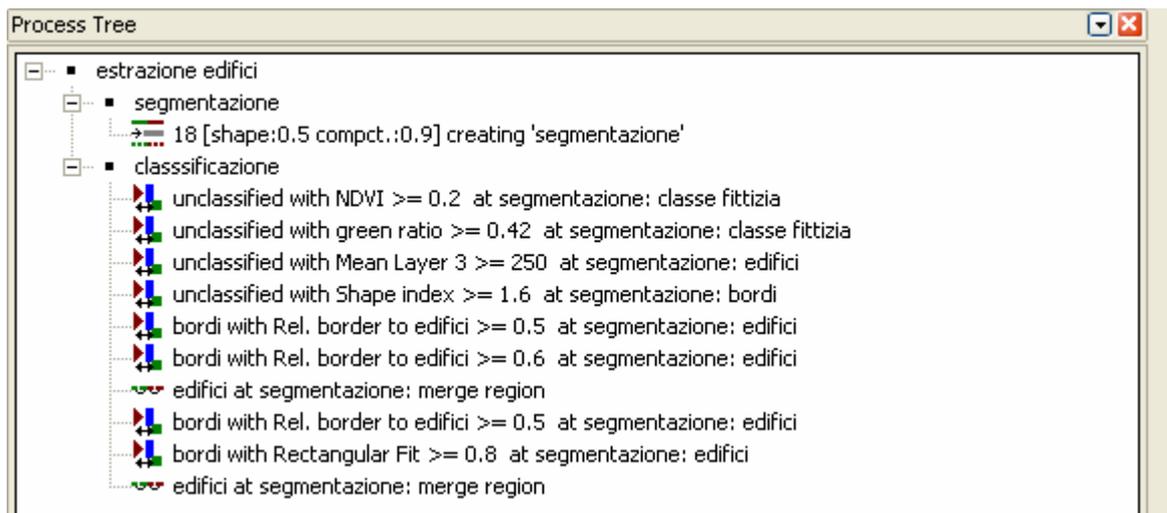
La zona di studio riguarda un'area in prossimità di Caselle (Torino), località famosa per la presenza dell'aeroporto. Il data set anche in questo caso è acquisito dal satellite QuikBird, le immagini hanno subito il pansharpening.

L'obiettivo anche in questa circostanza è stato quello di individuare gli edifici; sono stati necessari svariati tentativi perché il dato in questione risultava più complesso nell'analisi a causa della presenza di un grado d'ombreggiatura elevato che rende non omogenea la coperture degli edifici, nei quali le falde del tetto assumono un comportamento spettrale differente.



*In figura viene evidenziato il problema dell'ombreggiatura*

Un primo tentativo è stato sviluppato in analogia al caso di studio precedente. È stata effettuata segmentazione è del tipo multirisoluzione caratterizzata da un parametro scala pari a 18, shape 0.5 e compactness 0.9.

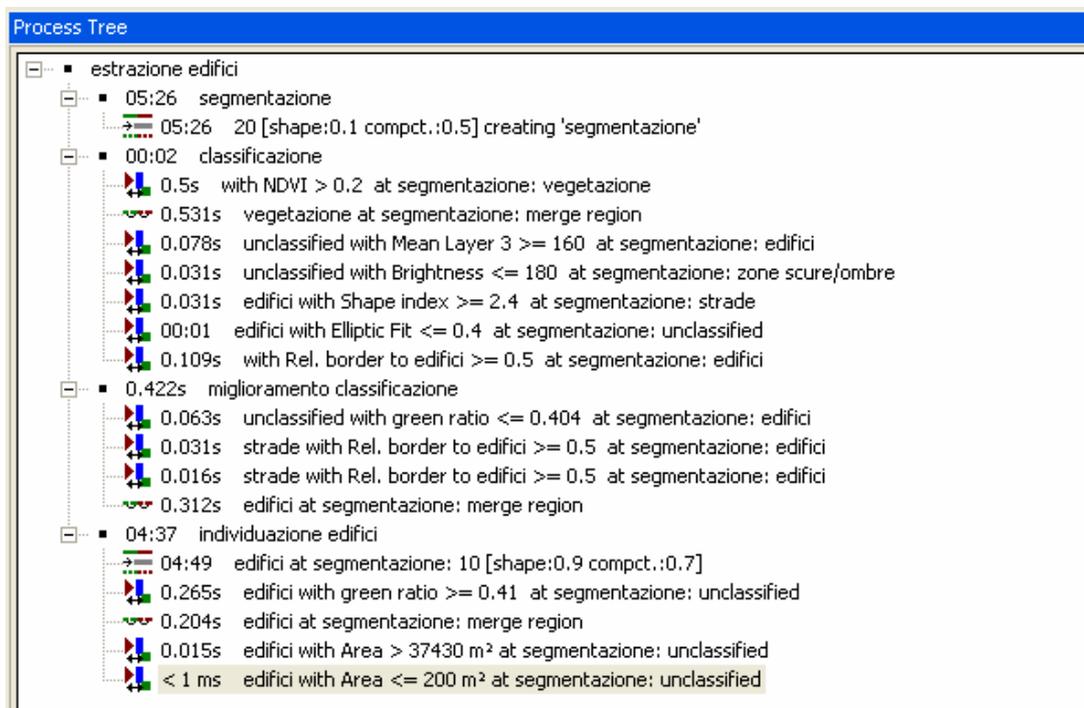


*Albero dei processi*

La classificazione è stata eseguita ricorrendo all'uso dell'algoritmo *assign class*. In questo caso viene utilizzata una soglia pari a 0.2 per l'indice di vegetazione NDVI. La classe edifici viene individuata ponendo un valore di soglia della banda del rosso  $> 250$ . Purtroppo i risultati non sono soddisfacenti quanto il caso precedente relativo a Bologna. Nell'immagine a seguire sono campiti in rosa gli oggetti che il software tramite il processo di classificazione è riuscito ad identificare. Si nota come non siano di elevata qualità è come le falde del tetto in ombra non siano state classificate correttamente.



Dati i risultati ottenuti da questo processo, si è deciso di cambiare approccio al problema. È stata ancora una volta utilizzata una segmentazione del tipo multirisoluzione con un parametro di scala pari a 20, shape 0.1 e compactness 0.5. Per analizzare l'immagine sono state create delle classi fittizie per isolare al meglio la classe degli edifici; sono stati presi in esame con maggior riguardo le features: Shape Index, Elliptic Fit, per individuare objects che rientrano spettralmente nella classe edifici, ma che in realtà non ne fanno parte.



*Albero dei processi*

I risultati ottenuti da questo processo risultano migliori rispetto ai precedenti, rimangono tuttavia alcune coperture non classificate e alcune aree erroneamente classificate edifici.

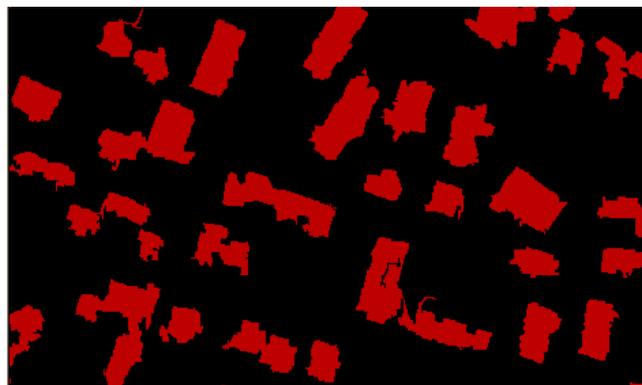


*Dettaglio dell'immagine classificata*

### **Valutazione dell'accuratezza**

Per questa caso di studio, relativamente al secondo risultato di classificazione, si è deciso di valutare l'accuratezza su di un subset dell'immagine.

Le statistica è stata prodotta col software Definiens Developer 7.0 e quindi in analogia ai casi precedenti è stato necessario creare l'immagine TTA Mask contenente gli edifici che rappresenteranno la *verità a terra*.



Porzione di TTA Mask utilizzata per la valutazione dell'accuratezza



*Subset dell'immagine utilizzato per la valutazione dell'accuratezza*

Il report in uscita ci fornisce un'accuratezza globale pari allo 0.78, valore probabilmente dovuto alla lacuna nella classificazione di identificare le coperture più chiare.

Error Matrix based on TTA Mask		
User \ Reference Class	EDIFICI	Sum
<b>Confusion Matrix</b>		
EDIFICI	100342	100342
unclassified	28318	28318
Sum	128660	
<b>Accuracy</b>		
Producer	0.78	
User	1	
Hitden	0.8763417	
Short	0.78	
KIA Per Class	0	
<b>Totals</b>		
<b>Overall Accuracy</b>	<b>0.78</b>	
<b>KIA</b>	<b>0</b>	

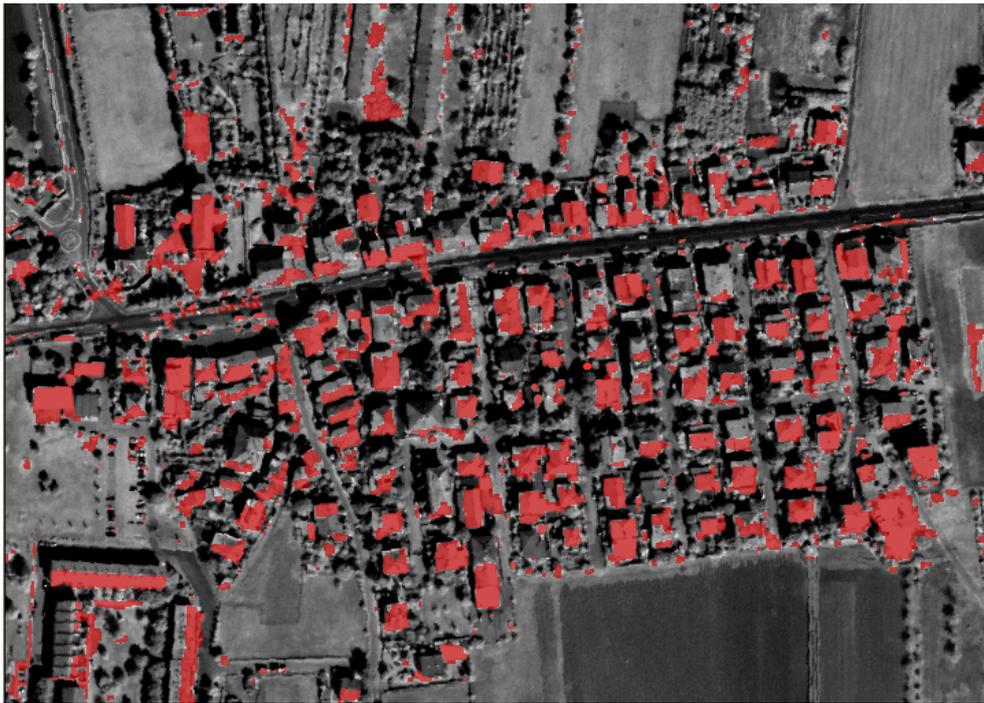
Se l'obiettivo fosse invece quello di individuare le coperture in laterizio, la classificazione risulta avere un'accuratezza pari allo 0.887.

User \ Referenc...	EDIFICI	Sum
<b>Confusion Matrix</b>		
EDIFICI	98454	98454
unclassified	12563	12563
Sum	111017	
<b>Accuracy</b>		
Producer	0.8868372	
User	1	
Hellden	0.94	
Short	0.8868372	
KIA Per Class	0	
<b>Totals</b>		
<b>Overall Accuracy</b>	<b>0.8868372</b>	
<b>KIA</b>	<b>0</b>	

### 5.2.3 Terzo caso di studio: Comune di Rimini

La zona di studio riguarda un'area della città di Rimini caratterizzata da un insediamento urbano localizzato circondato da aree agricole. In questo caso si ha a disposizione soltanto un'immagine in pancromatico acquisita dal satellite QuikBird: l'individuazione degli edifici risulta ovviamente più complicata perché non si hanno a disposizione informazioni spettrali e si ripresenta il problema dell'ombreggiatura sulle falde dei tetti.

La segmentazione è stata effettuata con l'algoritmo di segmentazione multirisoluzione contraddistinto da un parametro di scala pari a 5, è stato scelto un valore basso perché scegliendone uno alto non venivano colti a sufficienza i dettagli. Per quanto detto, la classificazione produce una accuratezza inferiore ai casi precedenti; per un suo miglioramento sarebbe necessario introdurre relazioni tra objects contigui.



*Porzione dell'immagine classificata: anche in questa circostanza  
la diversa ombreggiatura delle falde ha peggiorato il risultato*



# Capitolo VI

## Conclusioni

Gli ultimi anni hanno visto un crescente interesse per le procedure automatiche di classificazione di immagini digitali prodotte da satelliti o da riprese aeree. Tali procedure consentono la classificazione di ampie estensioni di territorio consentendo una riduzione del dispendio di tempo per il loro rilievo, e risultano utili ai fini di studi di uso del suolo, sviluppo urbano, salvaguardia ambientale e gestione delle risorse naturali.

La tesi ha riguardato l'applicazione di metodologie diverse di classificazione di immagini su problematiche differenti per scala di indagine e caratteristiche degli oggetti da estrarre. In particolare, sono state valutate le tecniche object-based che sono di recente sviluppo e che si dimostrano molto interessanti in quanto consentono elevate capacità di analisi dell'immagine. Esse consentono all'utente un grosso potere decisionale mettendogli a disposizione molteplici strumenti per l'interpretazione dell'immagine. Il processo ha un livello di automazione forse inferiore rispetto alle tradizionali metodologie pixel-based, è l'utente stesso a dover scegliere i parametri per personalizzare le fasi del workflow. L'innovazione che risiede nel metodo object-based è la fase di segmentazione: con questo processo si vanno a creare degli elementi chiamati objects che sono sostanzialmente delle aggregazioni di pixel che hanno valori radiometrici omogenei. In quest'ottica non si andranno a classificare i pixel individualmente come nella tecnica pixel-based, ma si lavora in termini di objects. Un primo vantaggio è nella diminuzione di elementi da classificare, ma la vera risorsa consiste nelle informazioni aggiuntive che gli objects forniscono: oltre alle

informazioni radiometriche, sono caratterizzati da informazioni contestuali, come ad esempio le relazioni tra elementi vicini, ma anche forma, tessitura, dimensioni, ecc.

Un vantaggio di questo modo di operare è la riproducibilità dell'analisi dell'immagine perché tutti i processi possono essere replicati in un secondo momento noti i parametri utilizzati. La strategia sviluppata per un caso particolare può essere applicata a dati della stessa tipologia, con un vantaggio non indifferente in termini di tempo in particolare in quelle circostanze in cui è necessario distinguere parecchie classi come nel caso urbano.

L'adozione di metodologia object-based è possibile tramite l'uso di software specialistici in grado di effettuare tutte le analisi richieste dall'utente; nell'ambito della tesi sono stati studiati due programmi professionali che sono attualmente considerati al massimo livello su questa problematica. Questi software risultano di notevole complessità e ricorrono all'uso di sofisticati algoritmi che possono essere personalizzati tramite l'impostazione di parametri; è possibile esportare mediante file vettoriali gli objects classificati sottoforma di poligoni permettendo un'immediata integrazione dell'analisi di dati telerilevati in ambito GIS.

La parte sperimentale della tesi ha riguardato due ambiti applicativi:

- la classificazione di piccoli oggetti su immagini close-range a colori naturali;
- la classificazione di edifici su immagini satellitari pancromatiche o multispettrali.

La prima sperimentazione si è incentrata sul conteggio e catalogazione automatica di frammenti di reperti ceramici riguardanti la missione archeologica italo-turcomanna locata nei pressi del conoide alluvionale del fiume Murghab (Turkmenistan) (Cerasetti, 2009). L'obiettivo è stato quello di ottenere un conteggio automatico dei reperti per creare un'alternativa valida all'attuale metodologia di catalogazione manuale molto onerosa in termini di tempo, di

personale specializzato e di gestione. La classificazione di immagini close-range, tramite la metodologia object-based, non ha ovviamente la stessa precisione e attendibilità che si hanno con un conteggio manuale, ma ha prodotto risultati abbastanza soddisfacenti in particolare se relazionati al rapporto costi-benefici, perché consente di risparmiare notevolmente in termini di tempo e personale. Si può affermare quindi che i risultati derivati da questa sperimentazione possono essere considerati una valida alternativa al conteggio manuale.

La sperimentazione relativa all'individuazione ed estrazione di edifici in aree urbane su immagini QuickBird mostra alcune delle possibili difficoltà che possono presentarsi in uno studio di immagini. Nel caso in particolare si evince come un'elevata ombreggiatura può rendere complessa l'analisi dell'immagine, oppure come sia complesso ricavare informazioni da dati pancromatici.

Anche in questa casistica le tecniche di classificazione object-based risultano notevolmente migliori rispetto alle tecniche di classificazione tradizionale, in quanto questa nuova metodologia permette un'analisi delle immagini con attributi spettrali e contestuali. In particolare nei casi di studio in ambito urbano, che sono caratterizzati da una elevata eterogeneità e dettaglio del territorio, risultano molto utili le features riguardanti la tessitura, la forma, le dimensioni, le relazioni tra oggetti.



# Bibliografia

BAATZ M., SHÄPE A., “Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation”, *Angewandte Geographische Informationsverarbeitung XII*, Wichmann, Heidelberg, pp. 12-23, 2000.

BRIVIO P.A., LECHI G., ZILIOLI E., “*Principi e Metodi di Telerilevamento*”  
Città studi edizioni, 2006

CERASETTI B.: “Missione archeologica congiunta italo-turcomanna nel conoide alluvionale del fiume Murghab (Turkmenistan). Studio del fenomeno urbano nel Bronzo e nel Ferro: scavo archeologico e analisi territoriale”, Dip. Archeologia, Università di Bologna, 2009.

DARWISH A., LEUKERT K., REINHARDT W., “Image Segmentation for the Purpose Of Object-Based Classification”

FELIPPE B. ANTUNES A., LINGNAU C., CENTENO DA SILVA J., “Object Oriented Analysis and Semantic Network for High Resolution Image Classification”, *Anais XI SBSR*, Belo Horizonte, Brasil, INPE, p. 273-279, 2003.

HEROLD M., SCEPAN J., MÜLLER A., GÜNTHER S., “Object-oriented mapping and analysis of urban land use/cover using IKONOS data”, *Proceedings of 22nd EARSEL Symposium “Geoinformation for European-wide integration*, Prague, 2002

## **Bibliografia**

---

NEUBERT M., MEINEL G., “Evaluation of Segmentation Programs for High Resolution Remote Sensing Applications”

NIEMEYER I., CANTY M., “Pixel-based and object-oriented change detection analysis using high-resolution imagery”

NIEMEYER I., J. CANTY M., “Object-Oriented Post-Classification of Change Images”, *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology*, Manfred Ehlers, Editors, pp.100-108, 2002.

WEIH R. C., Jr., RIGGAN N. D., Jr. “A Comparison of Pixel-based versus Object-based Land Use/Land Cover Classification Methodologies”

WONG T. H., MANSOR S. B., MISpan M. R., AHMAD N., SULAIMAN W. N. A. “Feature Extraction Based on Object Oriented Analysis”

Definiens, eCognition User Guide

ITT, ENVI User Guide