Università degli Studi di Napoli "Parthenope"
Dipartimento di Scienze e Tecnologie

Corso di Telerilevamento

Lezione 9 – B

Classificazione supervisionata delle immagini

Parte II

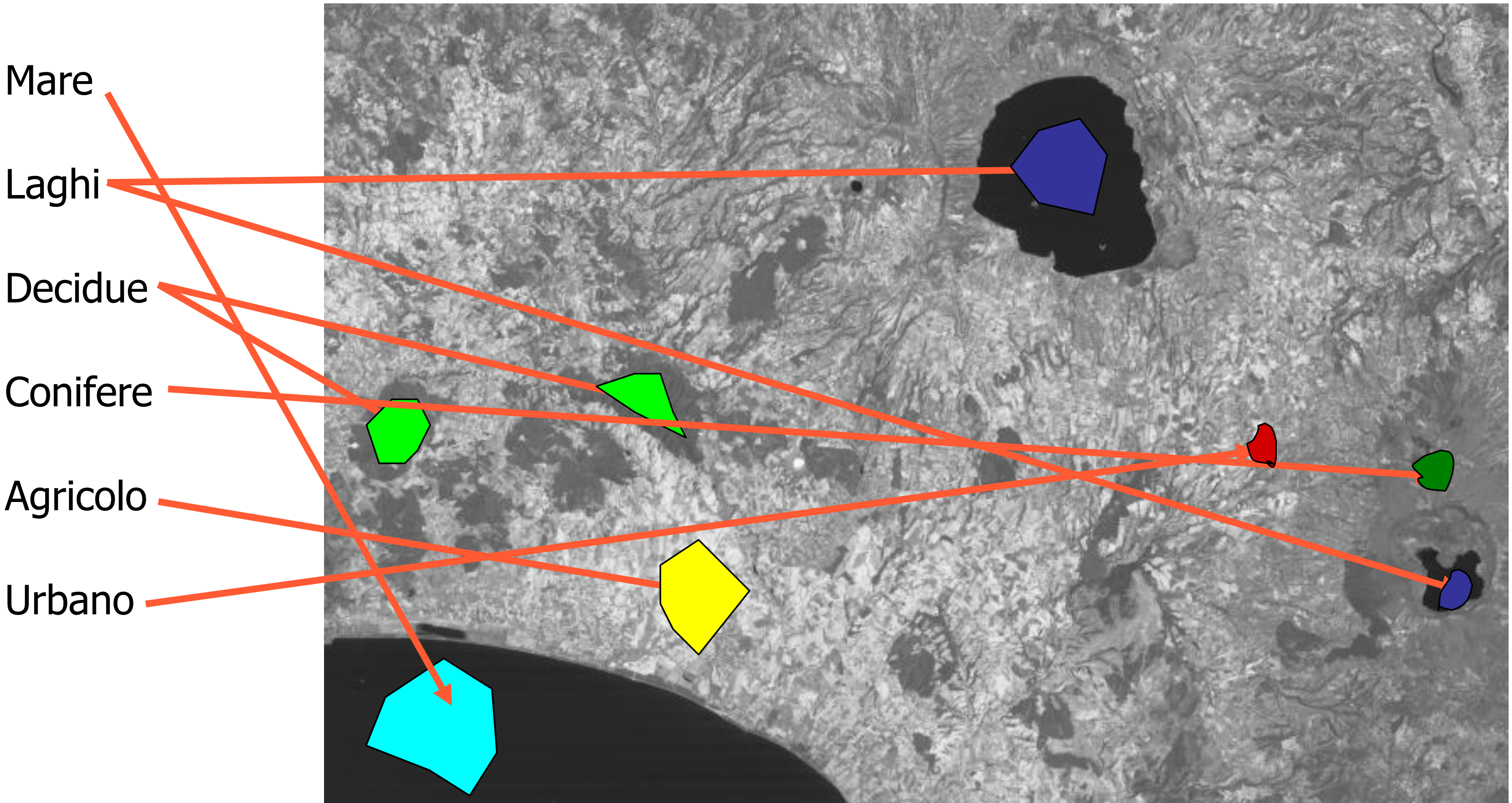
Claudio Parente

Classificazione supervisionata di una immagine singola

Ricordiamoci che stiamo sempre occupandoci di classificazione supervisionata e quindi ci serviamo di training sites per conoscere la caratterizzazione statistica delle classi in cui suddividere i pixels dell'immagine considerata.

Classificazione supervisionata di una immagine singola

Ricordiamoci che stiamo sempre occupandoci di classificazione supervisionata e quindi ci serviamo di training sites per conoscere la caratterizzazione statistica delle classi in cui suddividere i pixels dell'immagine considerata.



Classificatore della massima verosimiglianza – Immagine singola

Continuiamo ad occuparci della classificazione supervisionata applicata ad una semplice immagine.

Finora abbiamo visto il classificatore della minima distanza dalla media (MDM) e della minima distanza statistica.

Vediamo ora il classificatore della massima verosimiglianza.

Classificatore della massima verosimiglianza – Immagine singola

Probabilità condizionate = si applicano quando il numero dei pixel di ciascuna classe è approssimativamente lo stesso.

In caso di variazioni significative: modifiche dell'algoritmo di classificazione.

Esempio: numero dei pixel in una scena

n (vegetazione) = n (urbano) = 1/8 n (acqua).

Probabilità a priori p(i) = conoscenza a priori della probabilità che un pixel preso a caso appartenga ad una data classe.

Classificatore della massima verosimiglianza – Immagine singola

Probabilità a priori nell'esempio:

Vegetazione 0,1 (classe 2);

Urbano 0,1 (classe 3;

Acqua 0,8 (classe 1).

Per plottare le altezze dei relativi istogrammi non normalizzati si deve moltiplicare ciascuna curva degli istogrammi normalizzati per la rispettiva probabilità a priori.

Poiché l'interesse è per le altezze delle curve in modo relativo, basta moltiplicare la probabilità condizionata dell'acqua per 8, lasciando le altre invariate (Figura).

Classificatore della massima verosimiglianza – Immagine singola

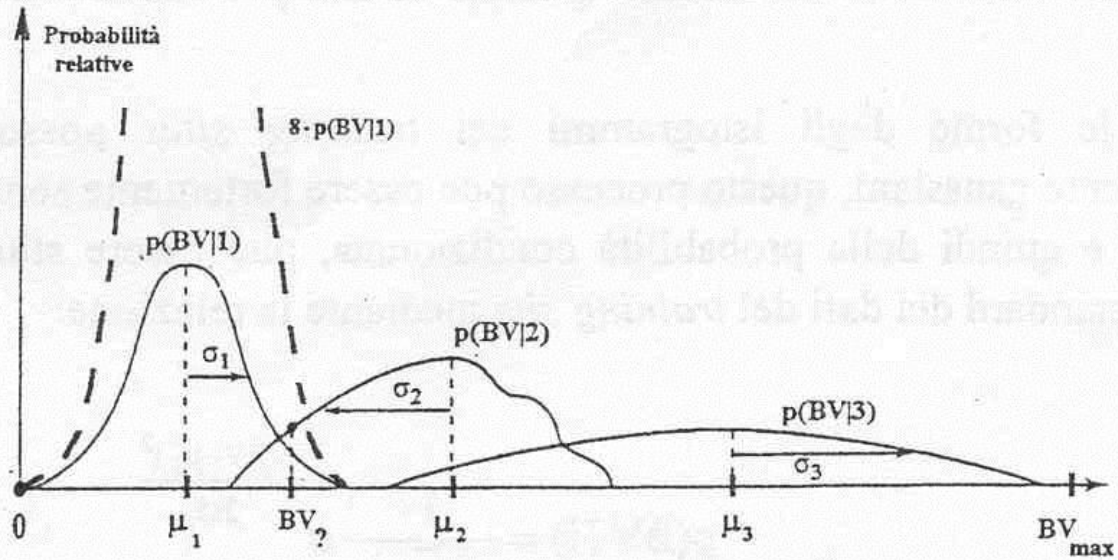


Fig.4.20. Effetto di una probabilità a priori su una classe. L'uso di probabilità a priori muterà la decisione sull'appartenenza di un $BV_?$ ad una determinata classe. In tal caso, molti più pixel aventi valori pari a $BV_?$ apparterranno alla classe 1 anziché 2 o 3.

il pixel BV? questa volta è molto più probabile appartenere alla classe 1 piuttosto che alla classe 2 o 3.

Classificatore della massima verosimiglianza – Immagine singola

Probabilità a priori = ***Teoria delle probabilità*** secondo ***Bayes***

Probabilità a posteriori che un pixel con un certo valore di BV appartenga alla classe i-sima:

$$p(i \mid BV) = \frac{P(BV \mid i) \cdot p(i)}{p(BV)}$$

dove p(BV) è la probabilità totale:

$$p(BV) = \sum_{i=1}^{l} p(BV \mid i) \cdot p(i)$$

Classificatore della massima verosimiglianza – Immagine singola

Il processo di classificazione = calcolo della probabilità a posteriori dei BV? assegnando i pixel con quel BV alla classe che fornisce la probabilità a posteriori più elevata.

In pratica p(BV) è la stessa per tutte le classi.

Le probabilità a priori possono essere ottenute:

o dalla conoscenza generale del contenuto della scena

o da una precedente mappa di classificazione della regione.

Classificatore della massima verosimiglianza – Immagine singola

Se non si ha alcuna informazione sulle probabilità a priori, esse possono semplicemente essere assunte uguali.

Le probabilità condizionate possono essere stimate dagli istogrammi normalizzati generati dai *training sites*.

Problema: enormi quantità di dati da memorizzare per generare una buona stima della vera classe.

Classificatore della massima verosimiglianza – Immagine singola

Caso unidimensionale = l'istogramma non è troppo largo

Esempio: immagine ad 8 bit contiene 256 valori per ogni classe.

Caso multispettrale: istogrammi sono multi-dimensionali e richiedono un'enorme quantità di dati per stimare accuratamente la loro forma.

Classificatore della massima verosimiglianza – Immagine singola

Se gli istogrammi dei *training sites* possono essere assunti approssimativamente gaussiani = processo semplificato.

La forma dell'istogramma, e quindi della probabilità condizionata, può essere stimata dalla media e dalla deviazione standard dei dati del *training site* mediante la relazione:

$$p(BV \mid i) = \frac{1}{\sqrt{2\pi}\sigma_i} \cdot e^{-\frac{(BV-\mu_i)^2}{2\sigma_i^2}}$$

Classificatore della massima verosimiglianza – Immagine singola

Combinando le equazioni e ignorando il termine P(BV), si ha

$$P_i^{'} = p(BV \mid i) \cdot p(i) = \frac{p(i)}{\sqrt{2\pi}\sigma_i} \cdot e^{-\frac{(BV-\mu_i)^2}{2\sigma_i^2}}$$

Per ciascun pixel vengono calcolati i valori dei P'_i nelle varie classi e il pixel viene assegnato alla classe con il valore di P'_i più elevato.

Si ottimizzerà così la probabilità di classificare correttamente il pixel sotto l'ipotesi di dati distribuiti secondo una gaussiana, da cui il nome per questo classificatore di massima verosimiglianza o GML (Gaussian Maximum Likelihhod).

Classificatore della massima verosimiglianza – Immagine singola

Ricorrendo alle espressioni logaritmiche, si ha

$$P_i'' = \ln P_i' = \ln[p(i)] - \frac{1}{2}\ln(2\pi) - \ln(\sigma_i) - \frac{(BV - \mu_i)^2}{2\sigma_i^2}$$

$$P_i = \ln P_i' + \frac{1}{2}\ln(2\pi) = \ln[p(i)] - \ln(\sigma_i) - \frac{(BV - \mu_i)^2}{2\sigma_i^2}$$

Classificatore della massima verosimiglianza – Immagine singola

Solo l'ultimo termine è funzione del BV ed è proporzionale al quadrato del numero di volte le deviazioni standard.

Pertanto, per una data classe, riducendo il numero di volte la deviazione standard, aumenta il valore di P_i .

Consideriamo ora il caso in cui le probabilità a priori riferite alle varie classi siano tutte eguali tra loro. In tal caso, detto k il numero delle classi, si ha:

$$p(i)=1/k.$$

Classificatore della massima verosimiglianza – Immagine singola

Assumendo le deviazioni standard relative ai training sites siano tutte uguali tra loro, indicando con σ il loro valore comune, si ha:

$$\ln(\sigma_i) = \ln(\sigma)$$

$$P_i = \ln\frac{1}{k} - \ln\sigma - \frac{(BV - \mu_i)^2}{2\sigma^2}$$

Classificatore della massima verosimiglianza – Immagine singola

Se si pone:

$$\ln\frac{1}{k} - \ln\sigma = \text{cost} = A_0$$

si ha:

$$P_i = A_0 - \frac{(BV - \mu_i)^2}{2\sigma^2}$$

Classificatore della massima verosimiglianza – Immagine singola

il valore di Pi è tanto più elevato quanto minore è la distanza euclidea |BV -μi|, ossia il problema si riconduce ad una classificazione MDM.

In altri termini, quando si prevede che le classi siano approssimativamente equi-distribuite ed i *training sites* rappresentati da istogrammi che presentano deviazioni standard paragonabili, può andar bene l'utilizzo di un *classificatore della minima distanza*. In altre condizioni, occorrerà utilizzare un classificatore della *massima verosimiglianza.*

Classificatore della massima verosimiglianza – Immagine singola

il classificatore GML può essere usato quando è possibile assumere che i dati si distribuiscono in modo "gaussiano". In caso contrario: funzione discriminante basata sulla massimizzazione delle probabilità a posteriori.

Ciò può essere relativamente semplice nel caso unidimensionale, ma operativamente sempre più difficile all'aumentare delle variabili, ossia delle bande spettrali.

Classificatore della massima verosimiglianza – Immagine singola

Di conseguenza, è spesso utile selezionare le classi le cui distribuzioni - in termini di BV - siano approssimativamente gaussiane. Ad esempio, una classe che si presenta bimodale, come quella riportata, potrebbe essere divisa in due sottoclassi unimodali di tipo gaussiano.

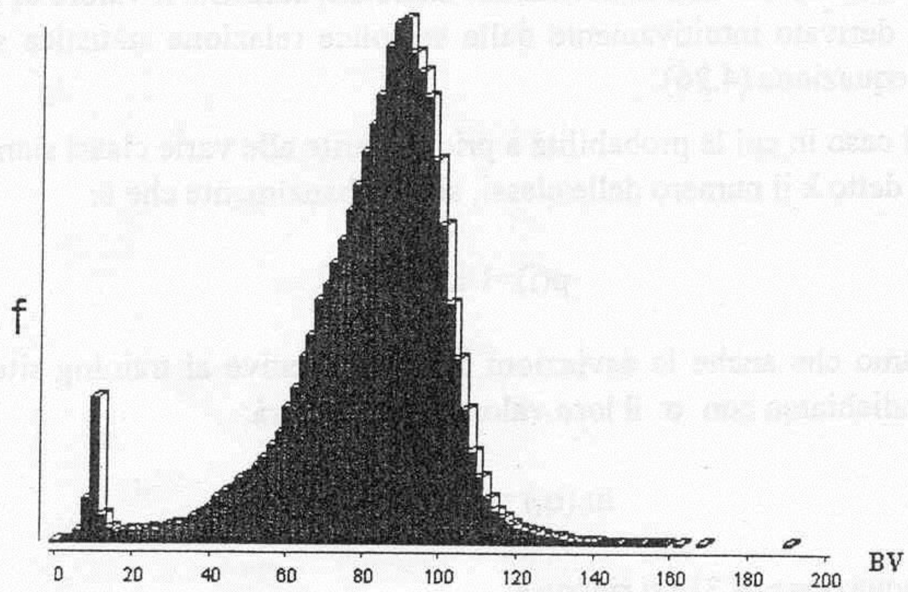


Fig.4.21. Esempio di un istogramma divisibile in due classi unimodali approssimativamente gaussiane centrate rispettivamente sui valori medi 10 e 90.

Classificazione supervised di immagini multispettrali

Il caso di classificazione di un 'immagine monocromatica è usato raramente (ad esempio, quando di una scena è disponibile solo una banda spettrale).

Comunque è base di partenza per estendere i concetti al caso di immagini multispettrali.

Si considerano immagini a due bande spettrali, anche se le soluzioni matematiche sono riferibili ad f bande spettrali.

Si consideri il set di *training sites* illustrato in figura dove sono plottate, con diversi simboli, tre classi per le due bande spettrali **banda** 1 e **banda 2.**

Classificazione supervised di immagini multispettrali

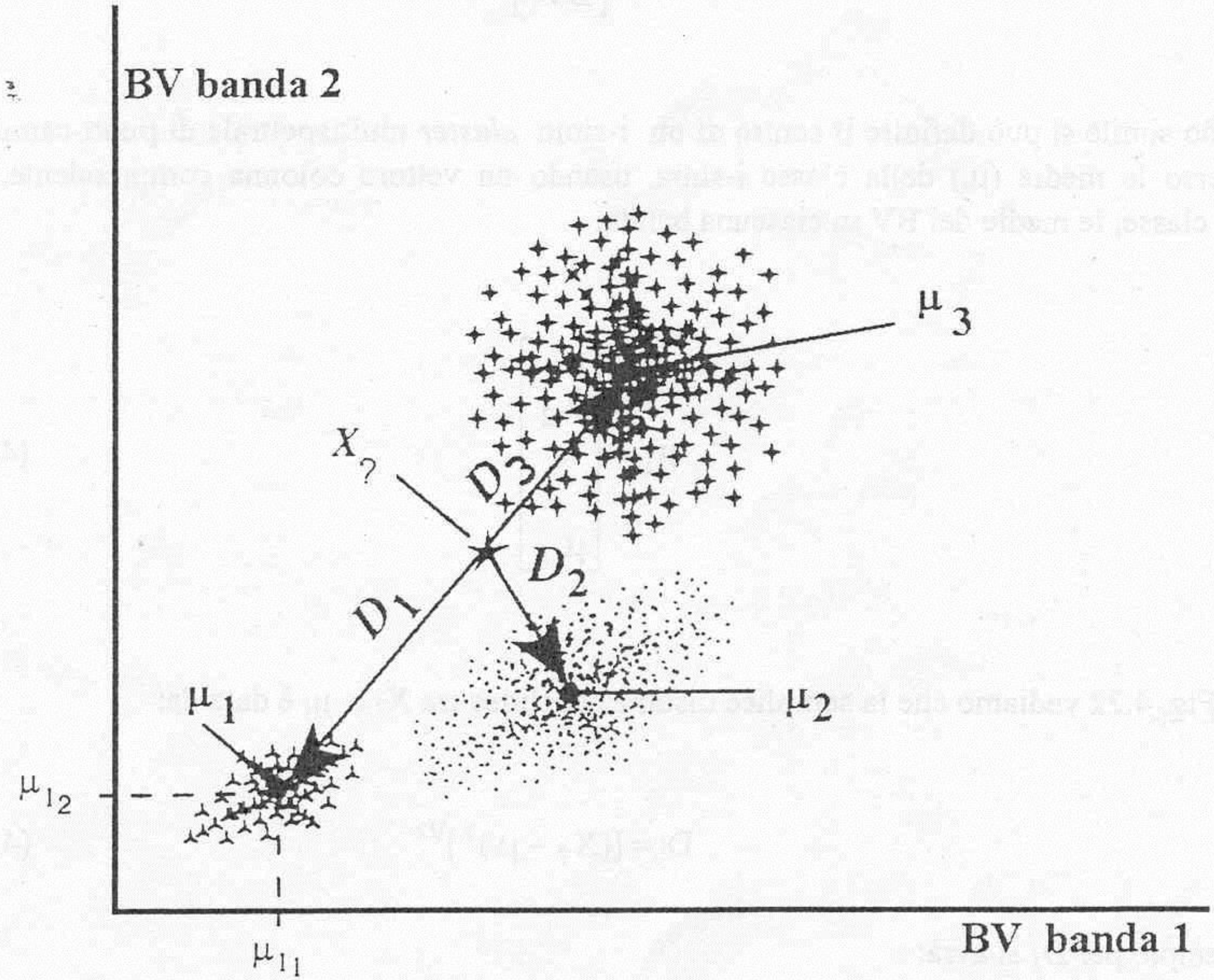


Fig. 4.22. Classificazione multispettrale usando un classificatore della minima distanza dalla media.

Classificazione supervised di immagini multispettrali

1 banda: i valori di BV proiettati su uno dei due assi = la discriminazione in considerevole sovrapposizione.

2 bande: gli scatterogrammi appaiono ben distinti, non essendovi alcuna sovrapposizione tra essi.

Obiettivo: definire una serie di algoritmi di classificazione che possano prendere vantaggio da questa apparente separabilità quando si usano più bande spettrali.

Classificatore multispettrale della minima distanza

Definiamo per ogni pixel un vettore colonna (X?) costituito dai suoi valori di BV
in ciascuna delle due bande spettrali cioè:

$$\mathbf{X}_? = \begin{bmatrix} BV_1 \\ BV_2 \\ \cdot \\ \cdot \\ BV_\ell \end{bmatrix}$$

Classificatore multispettrale della minima distanza

In modo simile si può definire il centro di un i-simo *cluster* multispettrale di punti-campioni .
attraverso la media (μ_i) della classe i-sima, usando un vettore colonna comprendente, per quella classe, le medie dei BV in ciascuna banda

$$\mu_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ . \\ . . \\ \mu_{i\ell} \end{bmatrix}$$

Classificatore multispettrale della minima distanza

La semplice distanza euclidea tra X? E μ_i è data da

$$D_i = [(X_? - \mu_i)^2]^{1/2}$$

$$D_1 = [(BV_1 - \mu_{11})^2 + (BV_2 - \mu_{12})^2]^{1/2}$$

Si può quindi implementare un classificatore della minima distanza (MDM) calcolando, per ogni pixel, il valore di Dj relativo alle *l* classi ed assegnando il pixel a quella classe che ha il valore minimo della distanza Di.

Classificatore a parallelepipedo

Classificatore MDM = molto semplice da implementare in un *software* molto veloce nell'esecuzione.

Ignora la variabilità dei dati.

Esempio: utilizzando il criterio della minima distanza, il pixel X? In figura verrebbe assegnato certamente alla classe 2.

Considerando la dispersione dei dati, esso va assegnato intuitivamente alla classe 3 che è caratterizzata da una maggiore varianza rispetto alla classe 2. In altre parole, X? appare più vicino al bordo del *cluster* 3 che a quello del *cluster* 2.

Classificatore a parallelepipedo

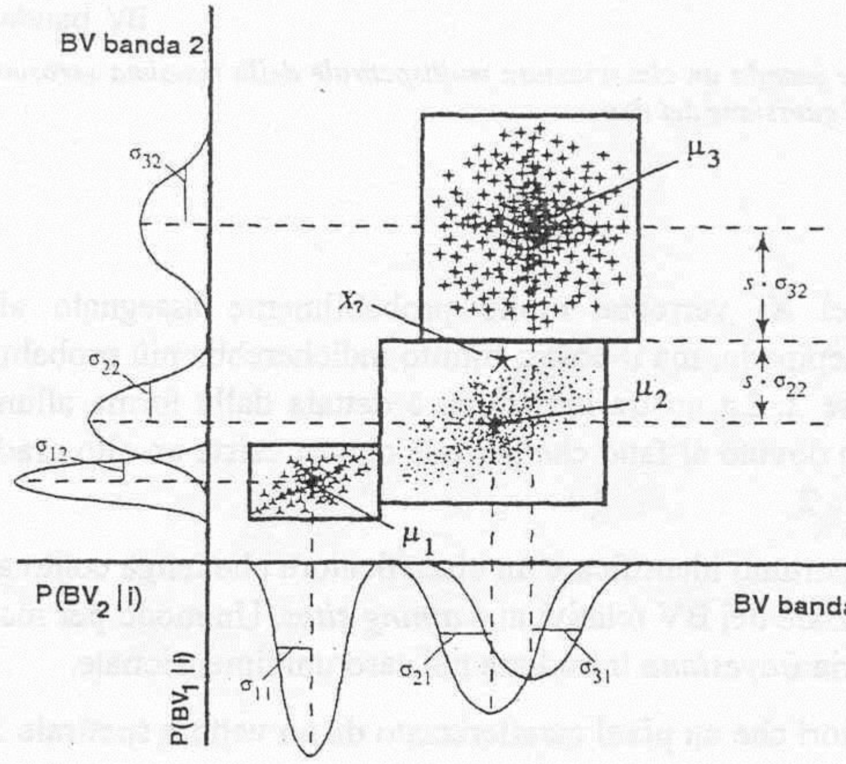


Fig.4.23. Classificatore a parallelepipedo usando per i limiti un ugual numero di deviazioni standard. Per localizzare il confine della classe 2 tra la classe 2 e la classe 3, viene equalizzato il numero di deviazioni standard ($s \cdot \sigma$) a partire dalla media di ogni classe. In figura il pixel $X_?$ verrebbe assegnato alla classe 2.

Classificatore a parallelepipedo

Nel caso unidimensionale si è creata la distanza standardizzata impiegando il numero di deviazioni standard per tenere conto della varianza.

Nel caso multibanda le varianze dei dati possono essere inserite in un opportuno classificatore mediante la localizzazione dei contorni nello spazio multidimensionale paragonabili a quelli che in uno spazio bidimensionale sarebbero dei rettangoli (figura).

Classificatore a parallelepipedo

Ogni pixel che cade all'interno del rettangolo associato con la i-sima classe viene assegnato a quella classe. Un pixel che non cade in nessun rettangolo viene considerato «non classificato».

Questo tipo di classificatore è detto a **parallelepipedo** che prende il nome dal fatto che in uno spazio a tre dimensioni i suddetti contorni fornirebbero dei parallelepipedi.

Classificatore della massima verosimiglianza - Immagini multispettrali

Il classificatore a parallelepipedo, pur essendo molto semplice e pur tenendo conto delle varianze, non tiene però conto della forma spettrale delle distribuzioni dei campioni

Classificatore della massima verosimiglianza - Immagini multispettrali

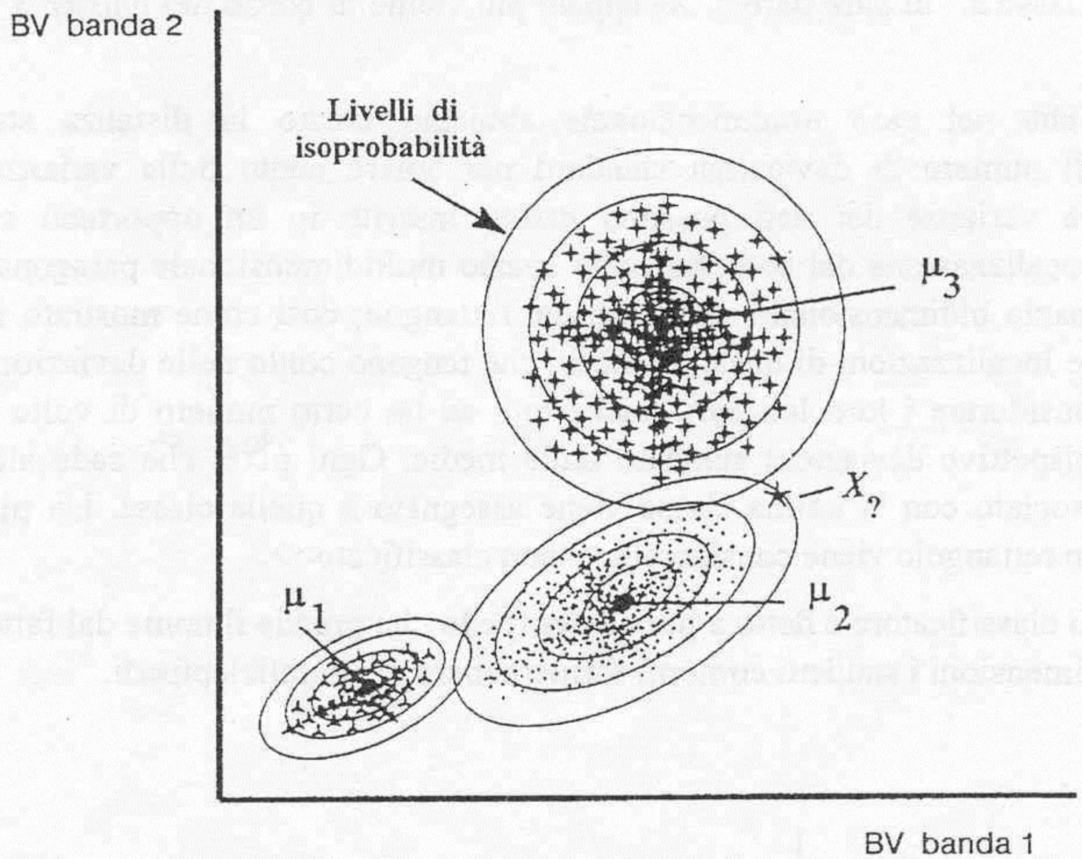


Fig. 4.24. Classificazione usando un classificatore multispettrale della massima verosimiglianza nell'ipotesi di distribuzioni gaussiane dei dati.

Classificatore della massima verosimiglianza - Immagini multispettrali

il pixel X? verrebbe molto probabilmente assegnato alla classe 3 dal classificatore a parallelepipedo, ma ad intuito parrebbe più probabile che il pixel sia un membro della classe 2.

L'intuizione è dettata dalla forma allungata dei dati del cluster 2, allungamento dovuto al fatto che per tale cluster esiste un alto grado di correlazione tra la banda 1 e la banda 2.

Classificatore della massima verosimiglianza - Immagini multispettrali

È opportuno identificare un classificatore che tenga conto anche della forma
della distribuzione spaziale dei BV relativi ai *training sites*. Un modo per includere tale forma è quello di usare la teoria *Bayesiana* introdotta nel caso unidimensionale.
La probabilità a posteriori che un pixel caratterizzato da un vettore spettrale X appartenga alla classe i-sima è:

$$p(i \mid X) = \frac{p(i) \cdot p(X \mid i)}{p(X)}$$