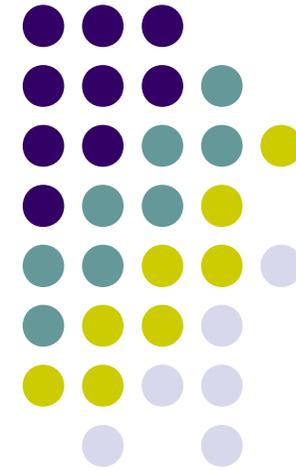


Dai pixel agli oggetti

Introduzione al riconoscimento di
oggetti
Funzioni di descrizione
Esempi



Riconoscimento di oggetti



- Come possiamo realizzare un sistema che automaticamente riconosca un certo tipo di oggetti in una scena ?
- Gli esseri umani sono estremamente abili per questo compito (Thanks, Mr. Darwin !)
- Estremamente difficile automatizzare tale capacità.



By the time they are five years old, most children can recognize digits and letters. Small characters, large characters, handwritten, machine printed, or rotated—all are easily recognized by the young. The characters may be written on a cluttered background, on crumpled paper or may even be partially occluded. We take this ability for granted until we face the task of teaching a machine how to do the same. Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make sound and reasonable decisions about the categories of the patterns. In spite of almost 50 years of research, design of a general purpose machine pattern recognizer remains an elusive goal.

A.K. Jain, R.P. Duin, J. Mao
“Statistical Pattern Recognition: a Review”

Riconoscimento automatico di oggetti (Pattern Recognition)



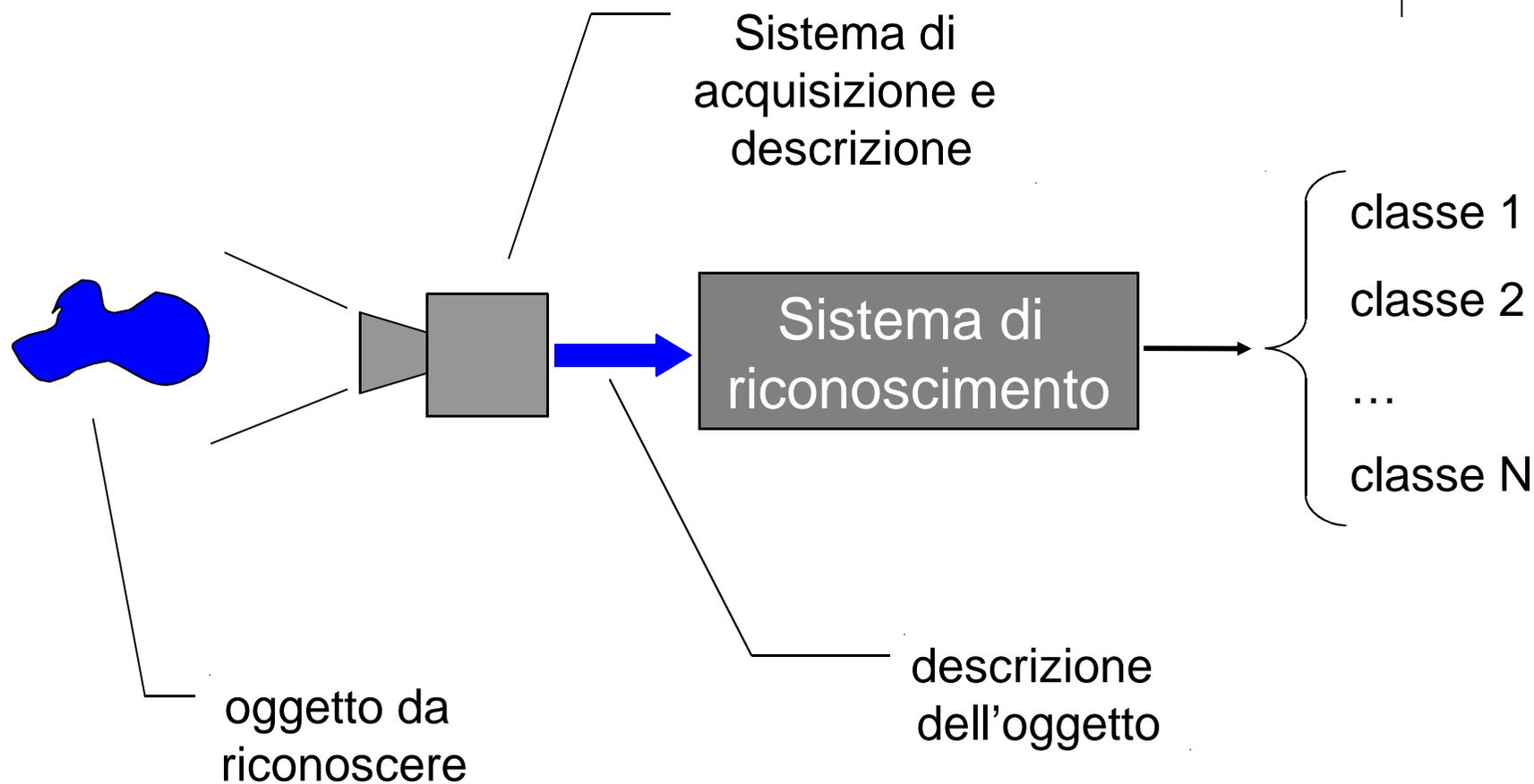
- Obiettivo possibile: definire un sistema per riconoscere automaticamente un oggetto
- data la descrizione di un oggetto che può appartenere ad una tra N classi possibili, compito del sistema è ascrivere l'oggetto ad una classe, utilizzando una base di conoscenza precedentemente costruita

Riconoscimento automatico di oggetti (Pattern Recognition)



- Obiettivo possibile: definire un sistema per riconoscere automaticamente un oggetto
- data la descrizione di un oggetto che può appartenere ad una tra N classi possibili, compito del sistema è ascrivere l'oggetto ad una classe, utilizzando una base di conoscenza precedentemente costruita

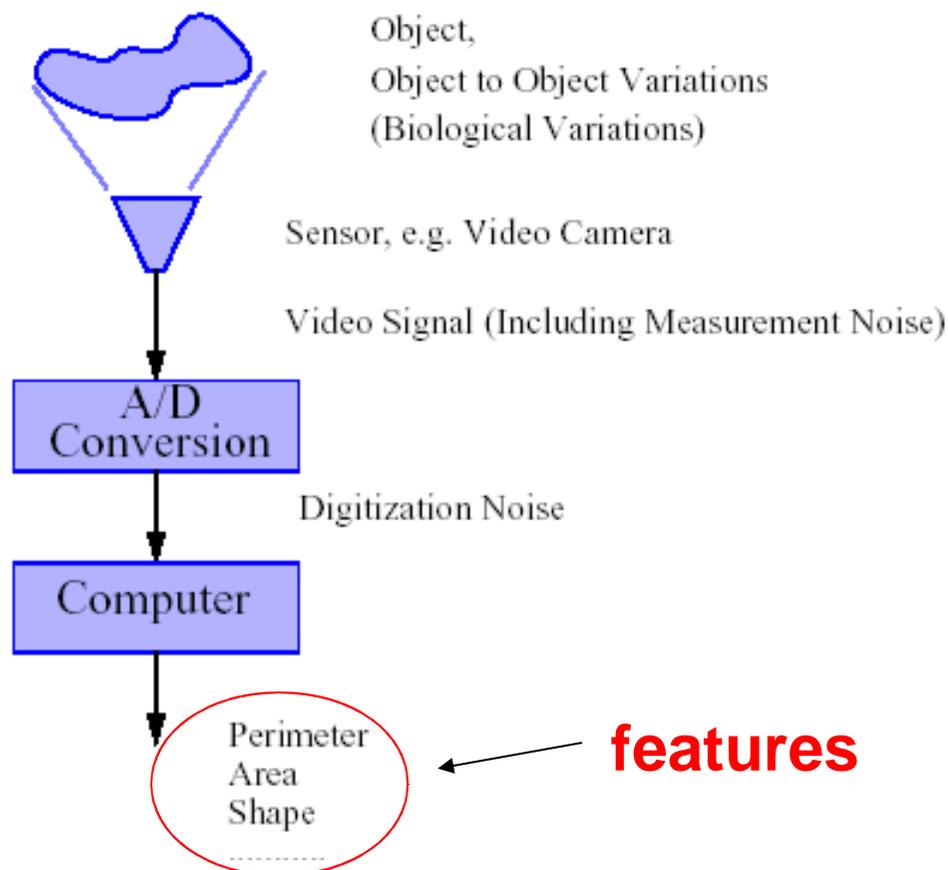
Schema ideale di un sistema di riconoscimento



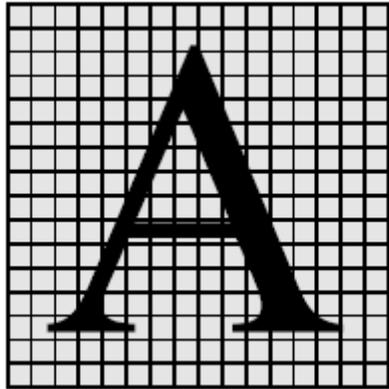


Descrizione

- In ingresso al sistema di riconoscimento è presentata una **descrizione**, cioè un insieme di misure (**features**) che **caratterizza** l'oggetto da riconoscere.
- L'insieme di misure è scelto sulla base delle esigenze specifiche

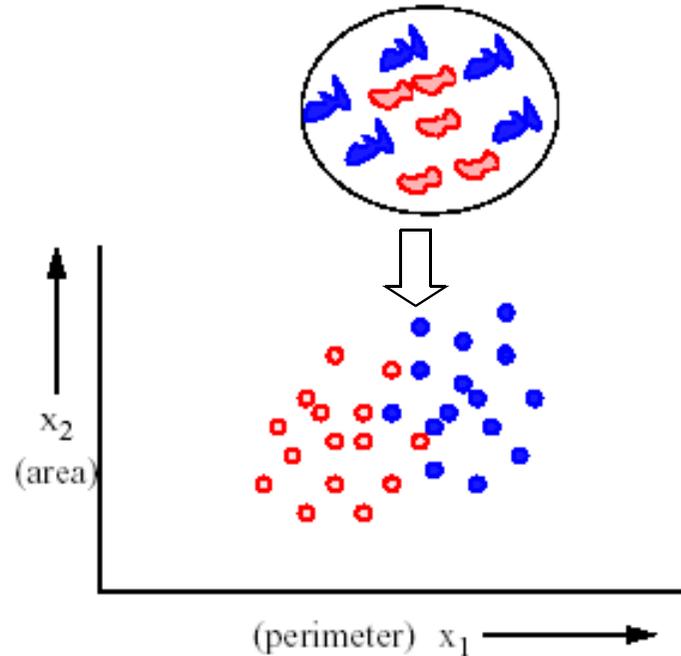


Rappresentazione e descrizione



- 16x16=256 pixel rappresentano l'oggetto
- da questi si possono ricavare un insieme di misure sintetiche (di solito in numero molto minore) che descrivono l'oggetto

Ad ogni oggetto corrisponde una descrizione in forma di vettore di features (*feature vector*) che corrisponde ad un punto in uno spazio delle features (*feature space*)



Problemi della rappresentazione in pixel



- L'input ad un classificatore è tipicamente un insieme di features che provengono dall'immagine, e non l'immagine stessa.
- Perché non possiamo usare direttamente i pixel ?
 - Gli oggetti corrispondono a regioni. E' necessario conoscere le relazioni spaziali tra i pixel.
 - In molti casi l'informazione non è nei pixel, ma nella forma.

Problemi della rappresentazione in pixel



- Inoltre, non è pensabile di realizzare un sistema efficiente di riconoscimento che si basa sul confronto diretto tra le matrici di pixel.
- La rappresentazione in pixel è infatti
 - ridondante
 - estremamente sensibile a modifiche anche insignificanti



Altri tipi di rappresentazione

- Oltre alla matrice di pixel, esistono altri tipi di rappresentazione più compatte e più utili ai fini del riconoscimento.
- Alcune di queste sono basate sul contorno dell'oggetto:
 - Chain code
 - Approssimazione poligonale
 - Signatures

Signatures

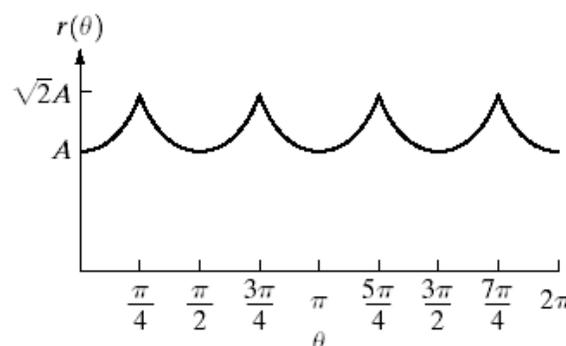
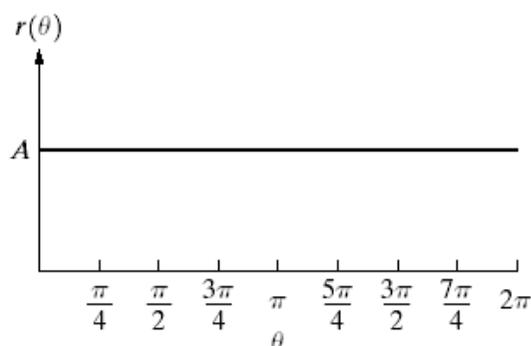
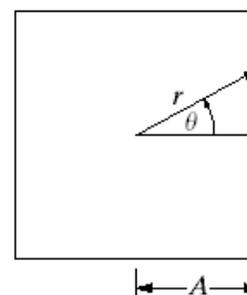
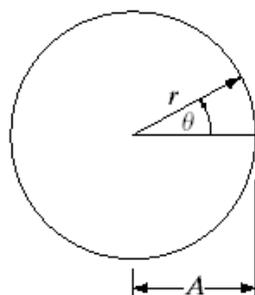


- Una *signature* è una rappresentazione monodimensionale del contorno di un oggetto e può essere valutata in vari modi.
- L'idea di base è di ridurre la rappresentazione del contorno (tipicamente bidimensionale) ad una funzione monodimensionale, presumibilmente più facile da gestire.



Signatures

- Un tipico esempio di signature è data dall'andamento della distanza dei punti del contorno dal baricentro dell'oggetto al variare di un angolo θ .



Signatures



- Questo tipo di rappresentazione è indipendente rispetto alla traslazione e può essere realizzata in modo da poter essere poco sensibile rispetto a
 - Rotazione
 - Scegliendo sempre lo stesso punto di inizio (es. il punto del contorno più lontano dal baricentro o l'intersezione con l'asse principale di inerzia)
 - Scala
 - Normalizzando i valori della funzione rispetto al valore massimo o alla deviazione standard

Dalla rappresentazione alla descrizione



Sebbene altri tipi di rappresentazione siano più efficaci rispetto alla semplice matrice di pixel, ai fini del riconoscimento risultano essere ancora ridondanti e sensibili a variazioni non significative.

Dalla rappresentazione alla descrizione



- E' quindi opportuno considerare, invece della *rappresentazione* dell'oggetto di interesse, una sua *descrizione*, cioè un insieme di misure o di proprietà qualitative che
 - permette di caratterizzare completamente l'oggetto ai fini del riconoscimento
 - è insensibile a variazioni non significative presenti sulle istanze dell'oggetto
 - consente di discriminare tra istanze di oggetti diversi

Caratteristiche della descrizione



- Non esiste una descrizione “buona per tutte le occasioni”
- La tipologia di descrizione va scelta in base
 - agli oggetti che si trattano
 - al contesto applicativo
 - al sistema di decisione
- Il discorso riguarda tutto l’ambito del Pattern Recognition, non solo quello dell’interpretazione delle immagini



Feature extraction

- La fase di scelta delle funzioni di descrizione (features) da utilizzare va sotto il nome di *feature extraction*.
- “Extracting from the raw data the information which is most relevant for classification purposes, in the sense of minimizing the within-class pattern variability while enhancing the between-class variability” (Devijver and Kittler, 1982)

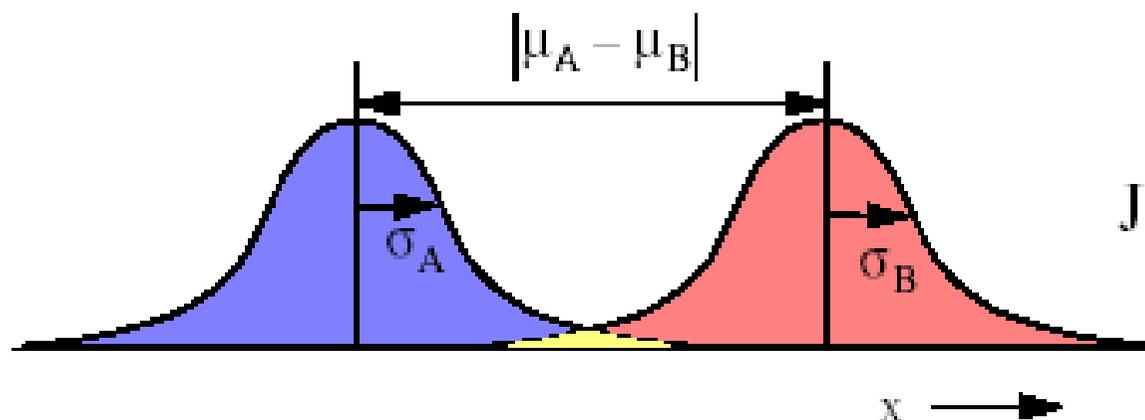


Feature extraction

Esistono diversi criteri per valutare il grado di discriminazione di un insieme di feature:

- Criteri euristici
- Criteri statistici

e.g. Fisher Criterion:



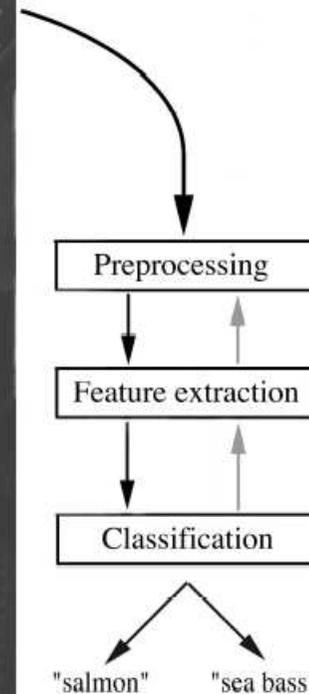
$$J_F(x) = \frac{|\mu_A - \mu_B|^2}{\sigma_A^2 + \sigma_B^2}$$

Un esempio ittico (Duda, Hart, Stork)



Problema:

distinguere il tipo di pesce (salmone o spigola) che transita su un nastro trasportatore

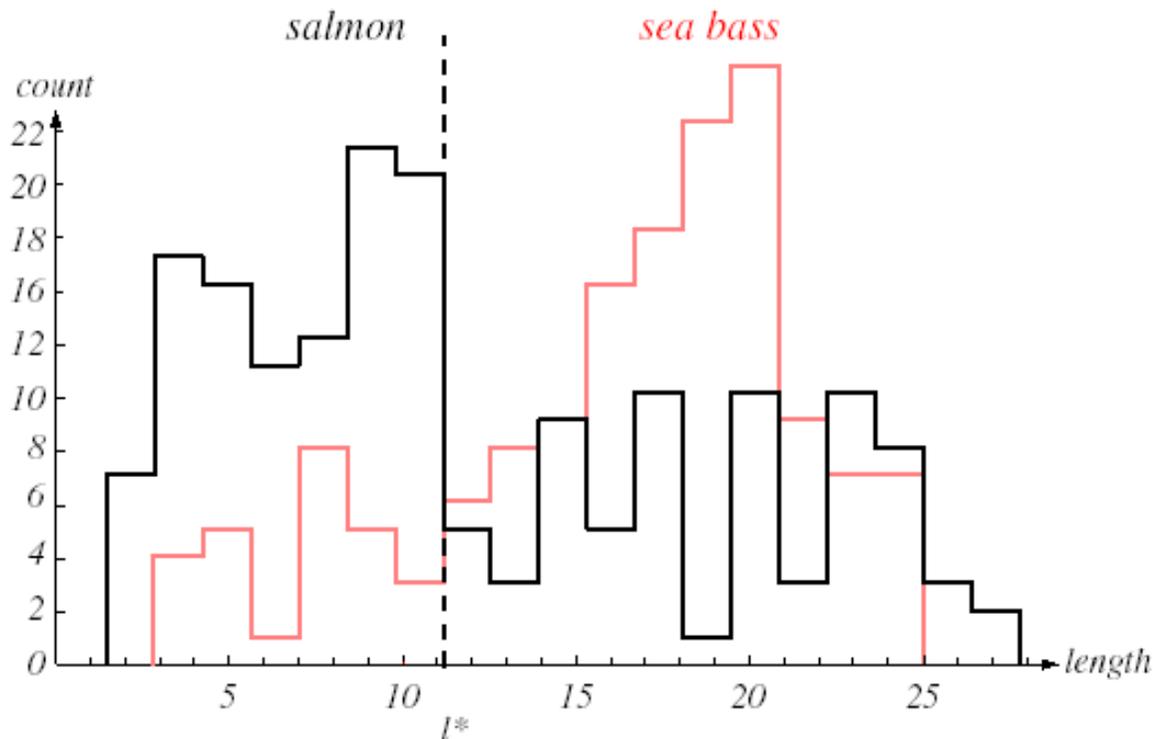




Scelta delle features

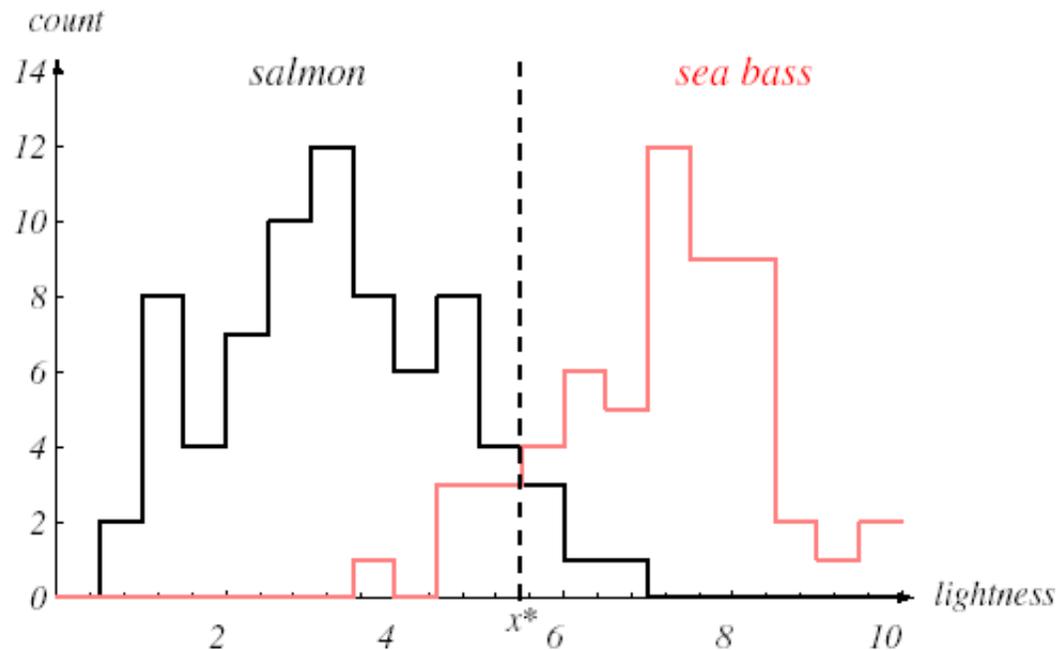
- Sulla base di informazioni ottenute sul problema si sceglie come feature iniziale la lunghezza.
- Per valutare l'efficacia di tale scelta e come questa possa essere usata è necessario considerare un insieme di campioni su cui effettuare la misura e fare le valutazioni. Si parla di *insieme di addestramento (training set)*.

Scelta delle features



La feature scelta non è molto discriminante. C'è una certa differenza, in media, ma non tale da separare nettamente le due classi.

Scelta delle features



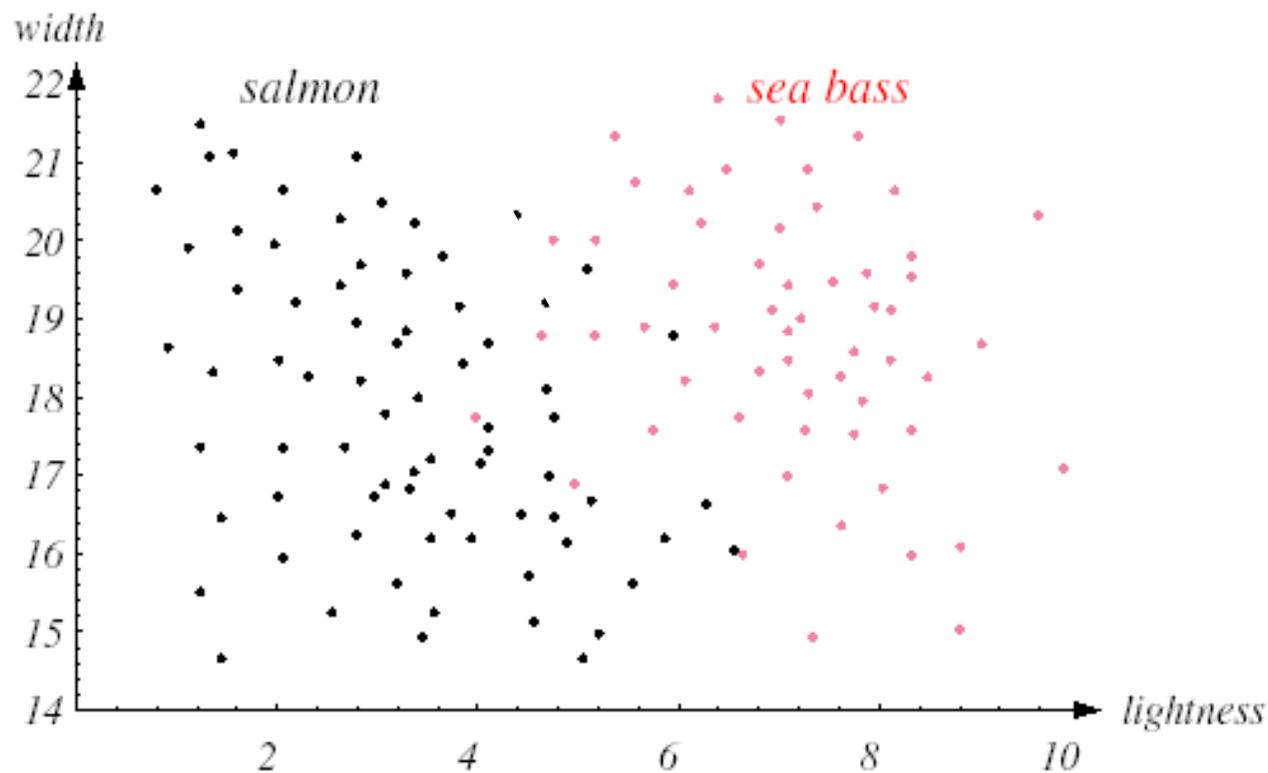
La nuova feature scelta permette una distinzione migliore tra le due classi. La soglia x^* è scelta in modo da minimizzare l'errore totale.



Scelta delle features

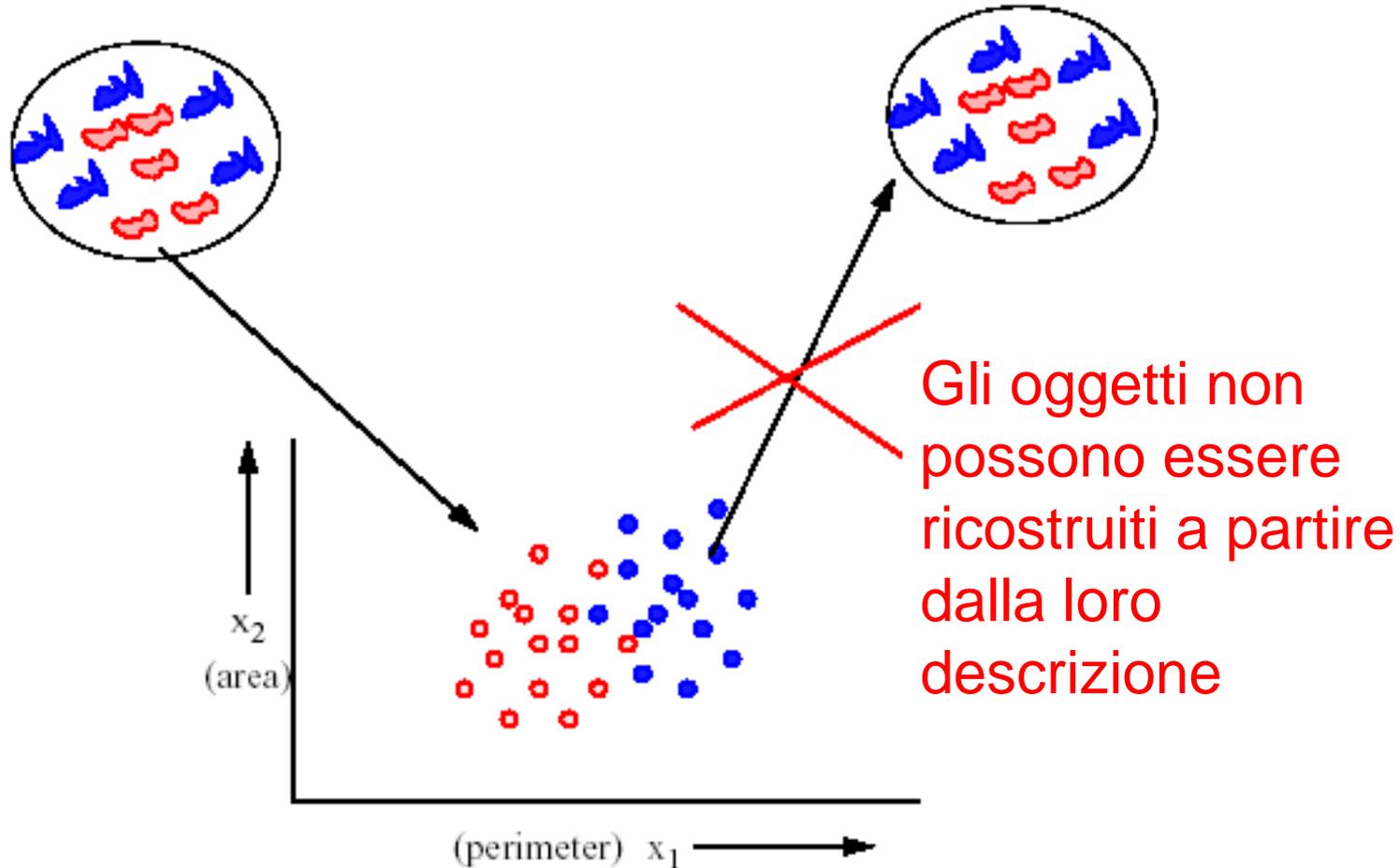
- La feature scelta, sebbene migliorativa rispetto al primo caso, potrebbe non portare ad una prestazione soddisfacente.
- Che cosa fare per migliorare ancora ? La soluzione più efficace è quella di scegliere un insieme di features da usare insieme.
- In questo caso, il singolo campione è descritto non più da un unico valore, ma da un insieme di valori.
- Formalmente ogni campione è descritto da un *vettore di features (feature vector)* a N componenti che individua un punto in uno spazio N-dimensionale (*spazio delle features o feature space*).

Spazio delle features



In questo caso si sono scelte due features. Lo spazio delle features è un piano ed ogni campione corrisponde a un feature vector $[x_1 \ x_2]^T$.

Differenza tra rappresentazione e descrizione



Tipologie di descrizioni



- Esistono due tipologie principali di descrittori:
 - Descrittori basati sul contorno
 - Descrittori globali (o basati sulla regione)

Esempi di descrittori basati sul contorno



Alcuni semplici descrittori di tipo geometrico basati sul contorno sono:

- **Lunghezza del contorno**

Viene valutata tramite il numero di attraversamenti tra pixel successivi compiuti attraversando il contorno dell'oggetto e contando 1 per gli attraversamenti orizzontali e verticali e $\sqrt{2}$ per quelli diagonali.

- **Diametro del contorno**

è definito come $\max_{i,j} d(P_i, P_j)$ dove P_i e P_j sono punti del contorno e $d(.,.)$ è una distanza sul piano digitale

Esempi di descrittori basati sul contorno

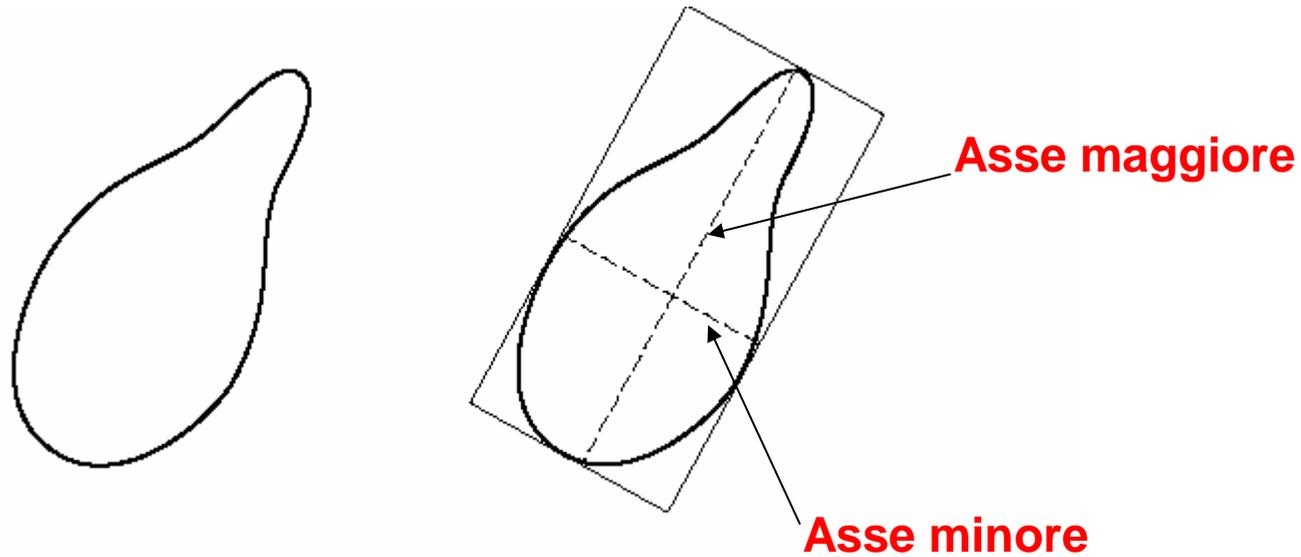


- **Asse maggiore, asse minore, eccentricità**

L'**asse maggiore** coincide con la linea su cui è stato valutato il diametro del contorno; l'**asse minore** è una linea perpendicolare all'asse maggiore e di lunghezza tale che il rettangolo passante per i quattro estremi dei due assi contiene completamente il contorno.

Il rapporto asse maggiore/asse minore viene definito **eccentricità** del contorno.

Esempi di descrittori basati sul contorno

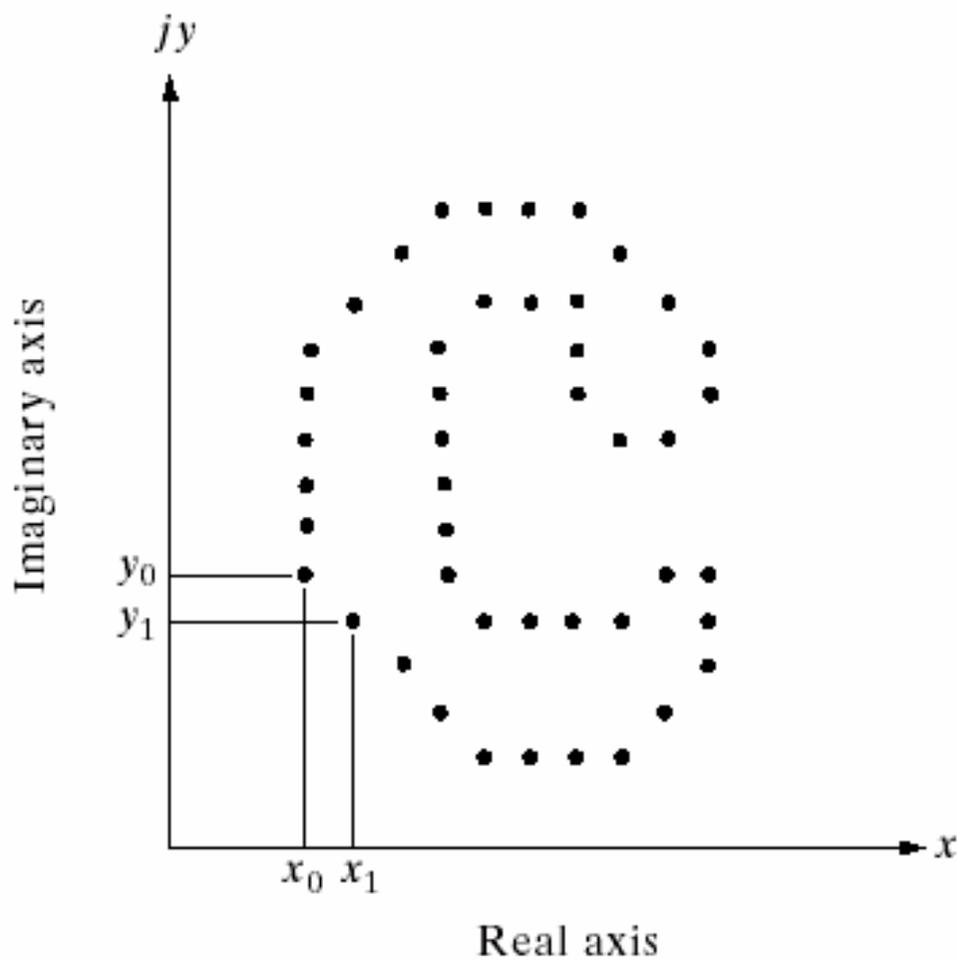


Descrittori basati sul contorno: descrittori di Fourier



- Dato il contorno di un oggetto e fissato un punto P_0 su di esso, si consideri la sequenza di punti $P_k = (x(k), y(k))$ che si ottiene attraversando il contorno a partire da P_0 .
- Se si assume l'asse x come asse reale e l'asse y come asse immaginario, la sequenza dei punti del contorno può essere assunta come sequenza di numeri complessi:

$$s(k) = x(k) + jy(k) \quad k=0, 1, \dots, K-1$$





Descrittori di Fourier

- Se si applica la trasformata discreta di Fourier (DFT) a $s(k)$ si ottiene:

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-j \frac{2\pi uk}{K}} \quad u = 0, 1, \dots, K-1$$

dove i coefficienti $a(u)$ si definiscono *descrittori di Fourier* del contorno.



Descrittori di Fourier

- Per riottenere la sequenza di partenza, si deve applicare la trasformazione inversa:

$$s(k) = \sum_{u=0}^{K-1} a(u) e^{j \frac{2\pi u k}{K}} \quad k = 0, 1, \dots, K-1$$

- Nel caso si considerino solo una parte dei coefficienti, si ottiene un'approssimazione di $s(k)$:

$$\hat{s}(k) = \sum_{u=0}^H a(u) e^{j \frac{2\pi u k}{K}}$$

$$k = 0, 1, \dots, K-1$$

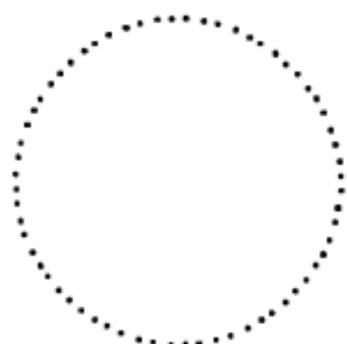


Descrittori di Fourier

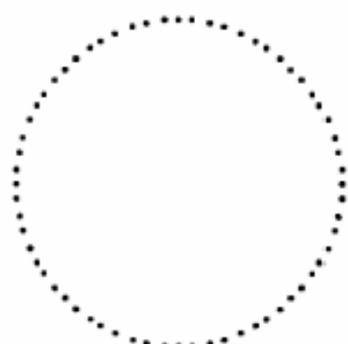
- Si noti che la sequenza approssimata contiene lo stesso numero di punti della sequenza originaria.
- La differenza tra le due è nel minor numero di dettagli presenti nella sequenza approssimata, dovuto al fatto di aver trascurato delle componenti ad alta frequenza nella ricostruzione.



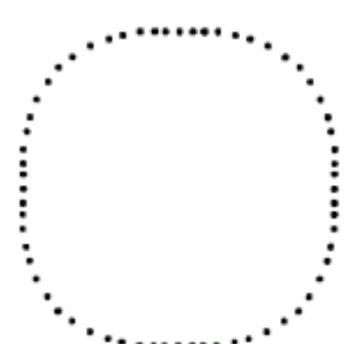
Original ($K = 64$)



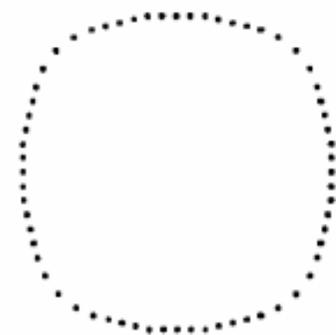
$P = 2$



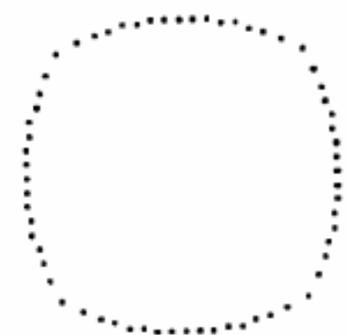
$P = 4$



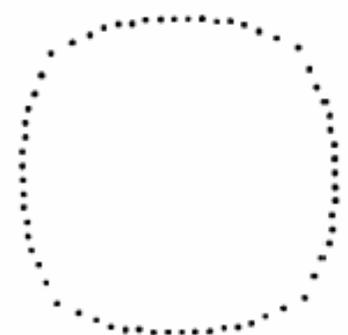
$P = 8$



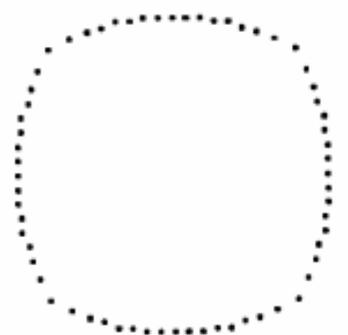
$P = 16$



$P = 24$



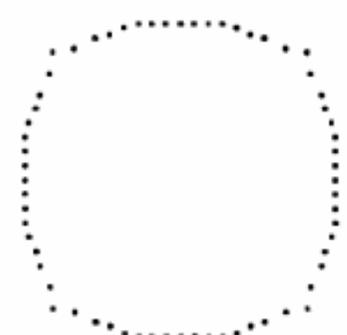
$P = 32$



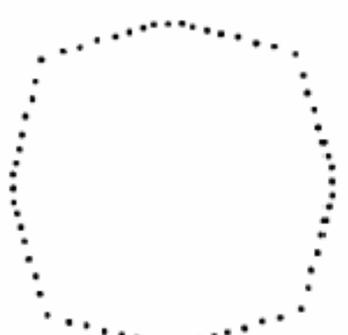
$P = 40$



$P = 48$



$P = 56$



$P = 61$



$P = 62$



Descrittori di Fourier

- Grazie alle proprietà della DFT è possibile predire quali siano gli effetti sui descrittori dei Fourier di traslazioni, rotazioni, variazioni di scala:

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

Esempi di descrittori globali



- **Area**

L'area di una regione si valuta tramite il numero di pixel appartenenti alla regione

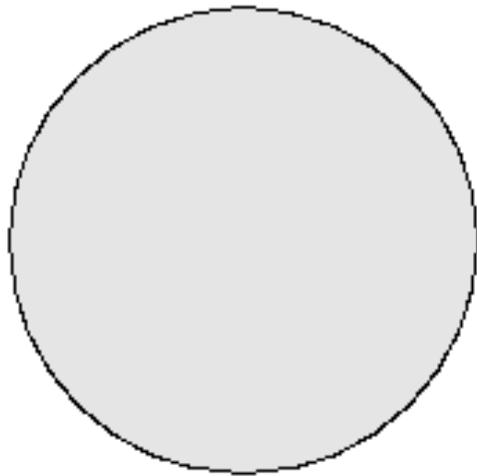
- **Perimetro**

Come perimetro di una regione si assume la lunghezza del contorno (vista prima)

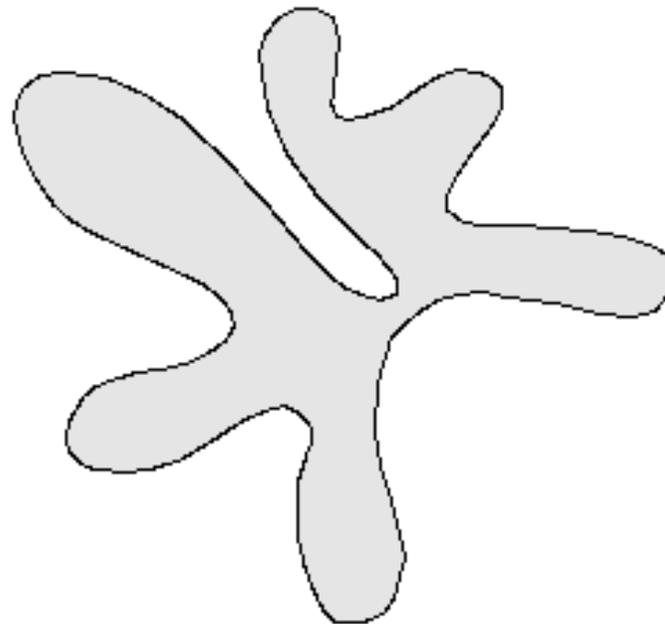
- **Compattezza**

è definita come $\frac{4\pi A}{P^2}$. Tende a 1 per forme vicine al cerchio

Esempi di descrittori globali: compattezza



Compattezza = 1



Compattezza $\ll 1$

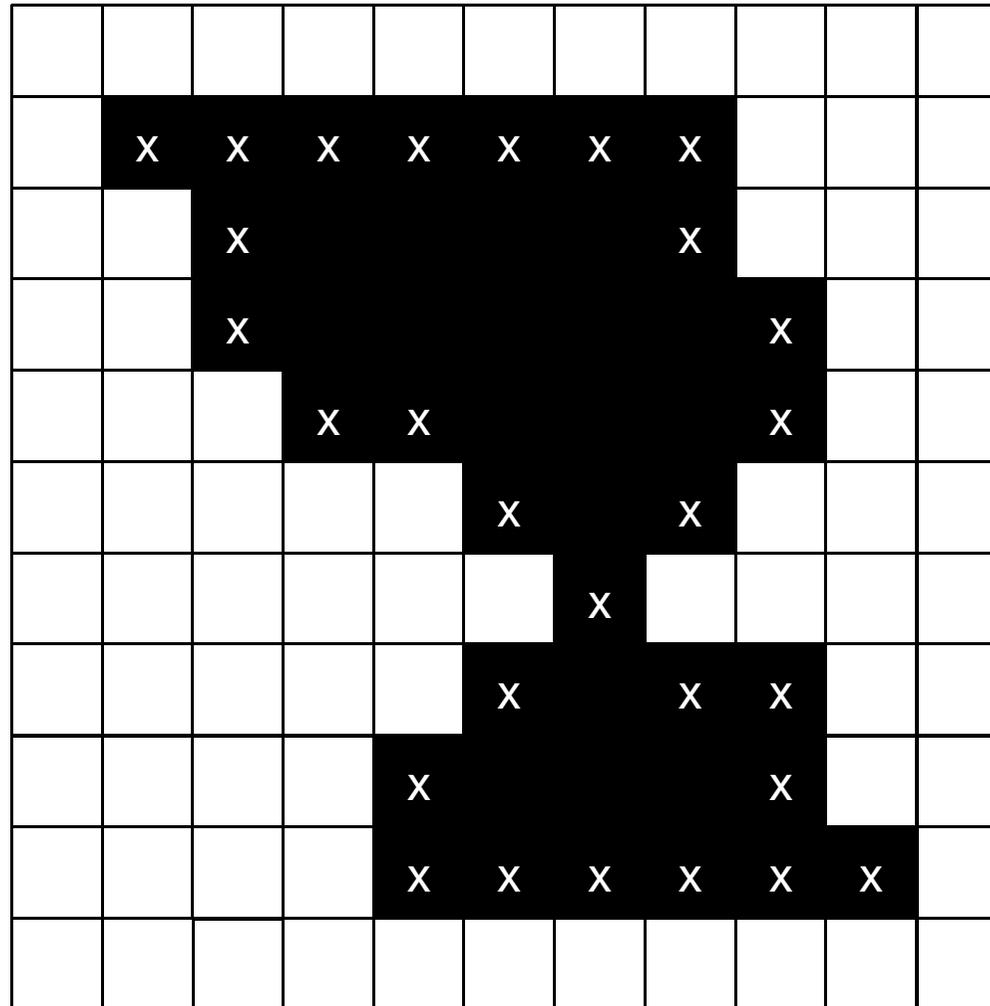
Esempi di descrittori globali: compattezza



$$P = 18 + 11\sqrt{2}$$

$$A = 41$$

$$\text{Compattezza} = 0.458$$

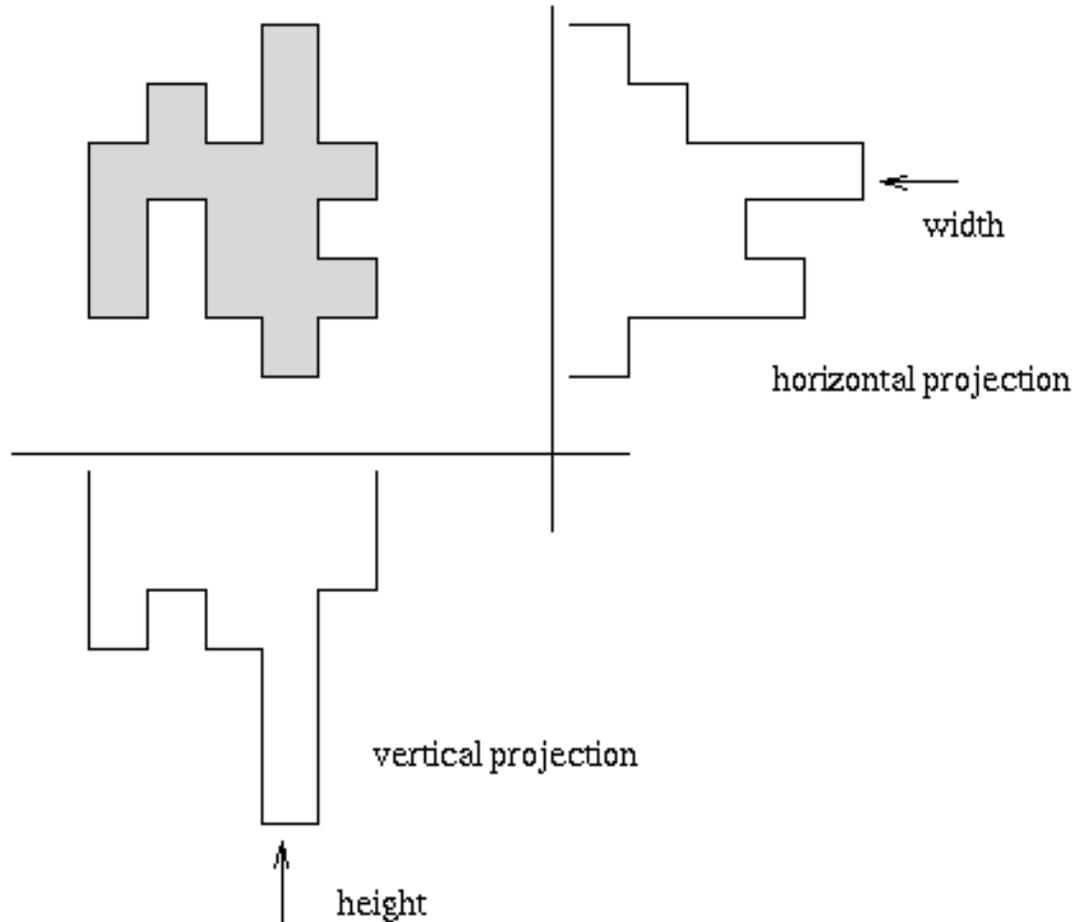


Esempi di descrittori globali: proiezioni



- Un'utile descrittore di tipo globale è fornito dalle *proiezioni*, che forniscono la distribuzione dei pixel della regione secondo alcune direzioni.
- La *proiezione verticale* è definita come il numero di pixel appartenenti alla regione in ogni colonna. La *proiezione orizzontale* è definita come il numero di pixel appartenenti alla regione in ogni riga. E' possibile definire anche *proiezioni diagonali*, che contano il numero di pixel appartenenti alla regione sulle diagonali.

Esempi di descrittori globali: proiezioni



Esempi di descrittori globali: proiezioni



- Oltre ad essere impiegate come descrittori globali, le proiezioni possono essere usate per separare regioni differenti. La presenza di un intervallo a valore nullo identifica, infatti, un gap tra regioni distinte.
- In una tecnica nota come *signature parsing* le proiezioni verticali vengono usate per separare regioni orizzontali distinte (es. righe di un testo); successivamente, si considerano, su ogni regione, le proiezioni orizzontali per separare regioni verticali distinte (es. caratteri su una riga di testo). Alternando queste due fasi in maniera ricorsiva è possibile identificare singoli oggetti distribuiti orizzontalmente e verticalmente in una regione ampia, come caratteri e disegni in una pagina di testo.

Esempi di descrittori globali: Momenti geometrici



- Una regione può essere descritta tramite una funzione a due variabili $f(x,y)$.
- Sia $f(x,y)$ una funzione che descriva quali punti appartengono alla regione di interesse:

$$f(x,y) = \begin{cases} 1 & \text{se } (x,y) \in \text{all'oggetto} \\ 0 & \text{se } (x,y) \notin \text{all'oggetto} \end{cases}$$

Esempi di descrittori globali: Momenti geometrici



- Il *momento geometrico* di ordine $(p+q)$ è definito come:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y)$$

- Al variare di p e q i momenti geometrici forniscono delle informazioni su come sono distribuiti i punti all'interno dell'oggetto.

Esempi di descrittori globali: Momenti geometrici



- Di particolare interesse sono i momenti di ordine inferiore
- Per $p=q=0$, si ottiene l'area dell'oggetto:

$$m_{00} = \sum_x \sum_y f(x,y)$$

Esempi di descrittori globali: Momenti geometrici



- Con i momenti di ordine 1 si individuano le coordinate del centroide o baricentro dell'oggetto.

$$m_{10} = \sum_x \sum_y x f(x,y)$$

$$m_{01} = \sum_x \sum_y y f(x,y)$$

$$x_c = \frac{m_{10}}{m_{00}}$$

$$y_c = \frac{m_{01}}{m_{00}}$$

Esempi di descrittori globali: Momenti geometrici



- Se si fissa l'origine nel centroide, si ottengono i momenti centrali, invarianti alla traslazione:

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x, y)$$

$$\mu_{00} = m_{00} \quad \mu_{10} = \mu_{01} = 0$$

Esempi di descrittori globali: Momenti geometrici



- Per ottenere dei momenti invarianti alla variazioni di scala, si considerano i *momenti centrali normalizzati*:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad \text{dove} \quad \gamma = \frac{p+q}{2} + 1$$

Esempi di descrittori globali: Momenti geometrici



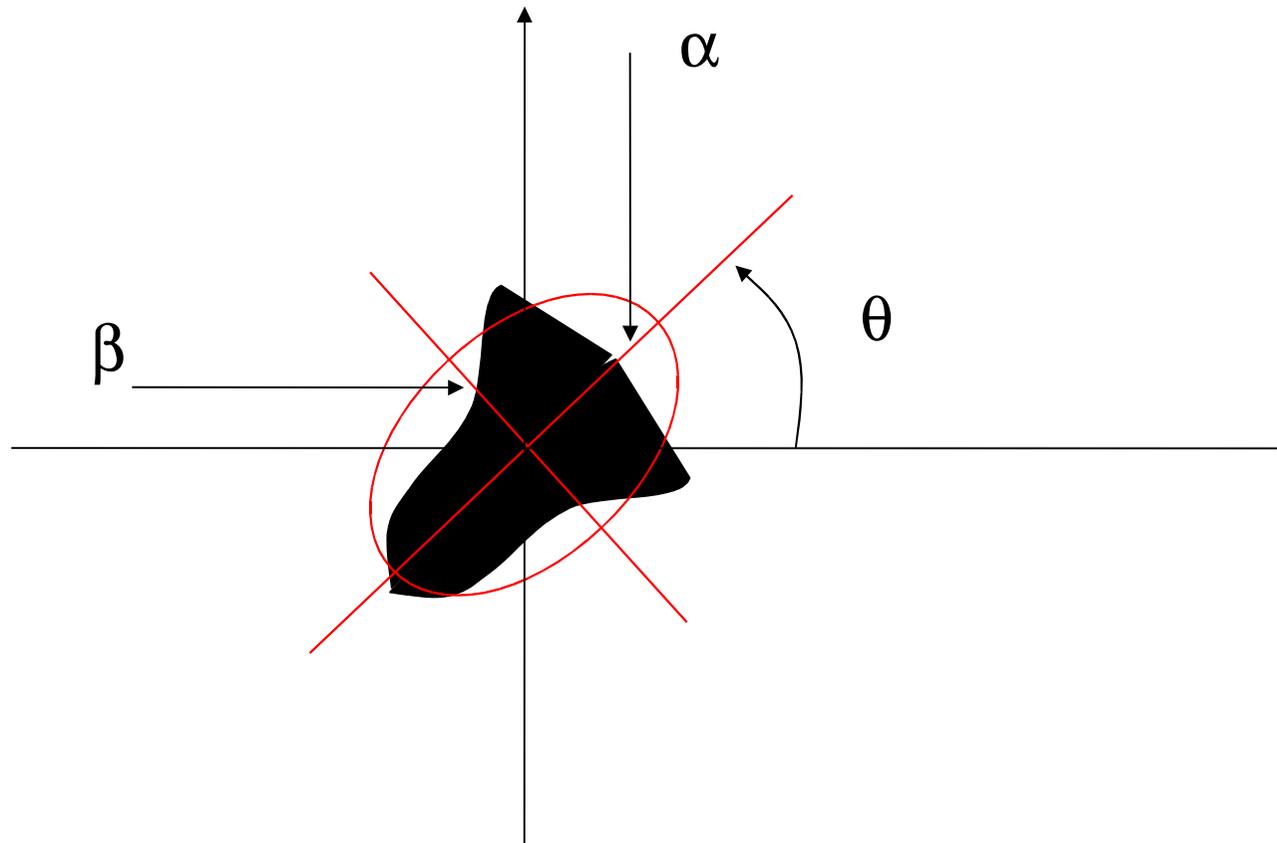
- I momenti del secondo ordine sono usati per determinare gli *assi principali d'inerzia* dell'oggetto.
- Questi forniscono utili indicazioni sull'orientazione dell'oggetto e sulla sua forma.

Esempi di descrittori globali: Momenti geometrici



- In particolare, i momenti del secondo ordine definiscono un'approssimazione dell'oggetto definita *image ellipse*.
- Questa è un'ellisse avente la stessa area, orientazione, eccentricità dell'oggetto ed è centrata in corrispondenza del suo centroide.

Momenti geometrici: *l'ellipse*





Momenti geometrici

- I parametri dell'immagine ellipse sono forniti da:

$$\alpha = \left(\frac{2 \left[\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}} \right]}{\mu_{00}} \right)^{1/2}$$

$$\beta = \left(\frac{2 \left[\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}} \right]}{\mu_{00}} \right)^{1/2}$$

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right)$$

Caratteristiche delle funzioni di descrizione



- Le funzioni viste sono *esempi* di possibili descrizioni per problemi di image understanding.
- A seconda delle esigenze del particolare problema, si possono considerare spazi delle features contenenti features di natura molto diversa che richiedono delle opportune trasformazioni.

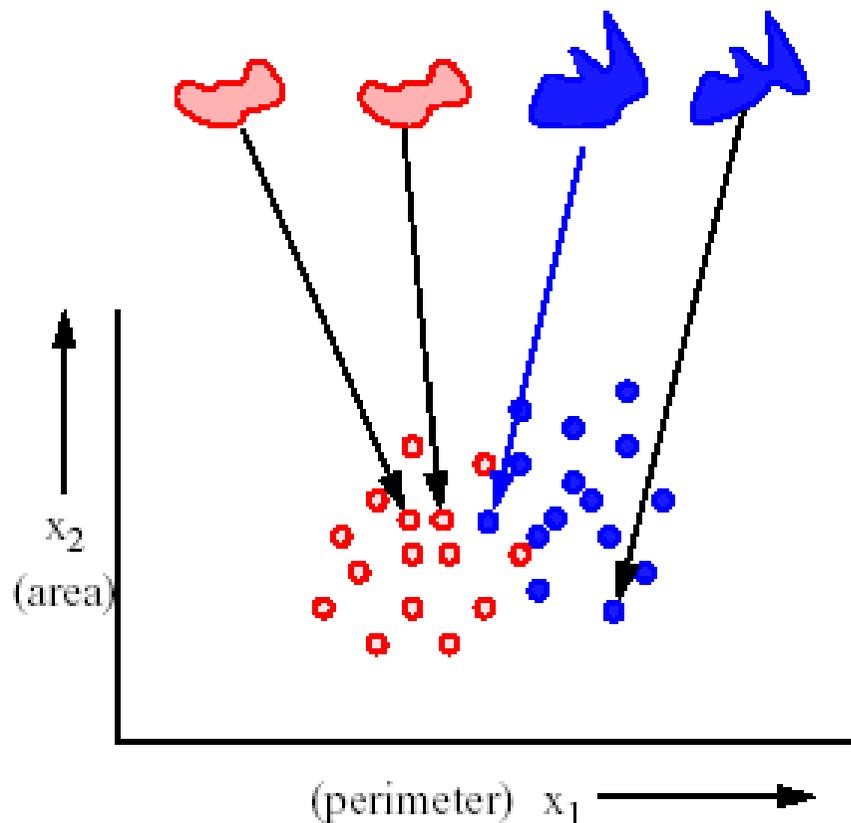


L'ipotesi di compattezza

Oggetti simili sono vicini nello spazio delle features

Achtung!

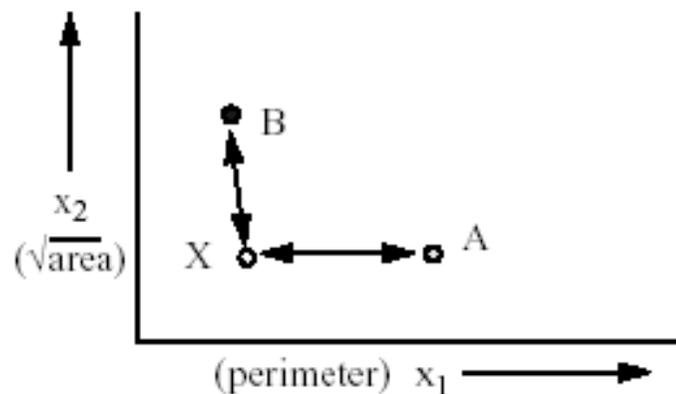
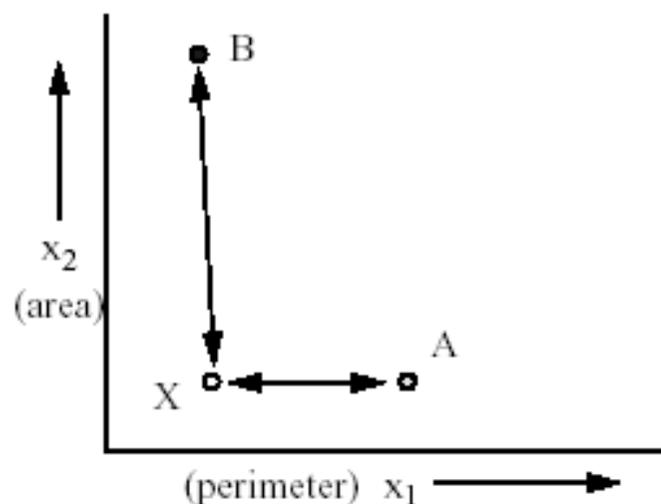
Oggetti diversi non sono necessariamente distanti



Normalizzazione delle features



In caso di features omogenee è opportuna una normalizzazione dimensionale



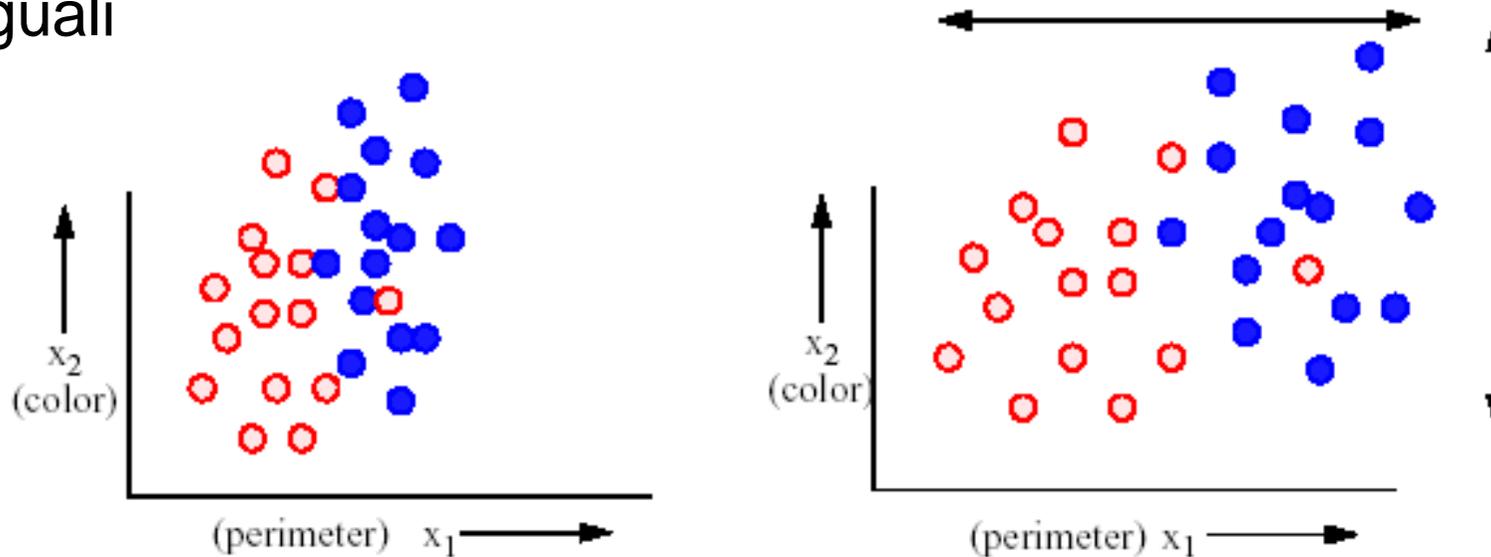
Prima della normalizzazione:
 $d(X,B) > d(X,A)$

Dopo la normalizzazione:
 $d(X,B) < d(X,A)$

Normalizzazione delle features



Anche se le features non sono omogenee potrebbe essere opportuna una normalizzazione che renda le varianze uguali



Prima della normalizzazione

Dopo la normalizzazione:

$$x' = \frac{x}{\sqrt{\text{var } x}}$$

$$y' = \frac{y}{\sqrt{\text{var } y}}$$