
Segmentazione di immagini

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Introduzione

Segmentazione: processo di partizionamento di un'immagine in regioni disgiunte e omogenee.



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Image segmentation



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik



Segmentazione

Consiste nel partizionamento di una immagine in regioni omogenee sulla base di un certo criterio di appartenenza dei pixel ad una regione

Obiettivo: individuare/riconoscere gli oggetti che compongono l'immagine

“The segmentation of the image(s) presented to an image analysis system is critically dependent on the scene to be sensed, the imaging geometry, configuration and sensor used to transduce the scene into a digital image, and ultimately the desired output (goal) of the system” (A. Jain)

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Regioni

Le regioni in cui l'immagine viene suddivisa devono soddisfare alcune proprietà:

- **DISTINCT**
 - Nessun pixel è condiviso da due regioni
- **COMPLETE**
 - Tutti i pixel dell' immagine sono assegnati ad almeno una regione della partizione
- **CONNECTED**
 - Tutti i pixel appartenenti ad una regione sono "connessi"
- **HOMOGENEOUS**
 - Tutte le regioni sono omogenee rispetto ad un criterio fissato (e.g. intensità, colore, texture, ecc..)

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Introduzione (def. formale) (1)

Sia R l'intera regione spaziale occupata dall'immagine. Il processo di segmentazione può essere visto come il partizionamento di R in n sottoregioni, R_1, R_2, \dots, R_n tali che:

$$\bigcup_{i=1}^n R_i = R.$$

R_i è un insieme connesso, $i = 1, 2, \dots, n$.

$R_i \cap R_j = \emptyset$ per tutti i valori i e j , $i \neq j$.

$Q(R_i) = \text{TRUE}$ per $i = 1, 2, \dots, n$.

$Q(R_i \cup R_j) = \text{FALSE}$ per ogni coppia di regioni adiacenti R_i, R_j .

Con $Q(R_k)$ predicato definito sui punti di un insieme R_k .

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Introduzione (def. formale) (2)

- Ogni pixel deve appartenere ad una regione;
- I punti appartenenti ad una regione devono essere connessi (es. 4-connessi, 8-connessi);
- Le regioni devono essere disgiunte;
- I pixel appartenenti ad una regione devono soddisfare un certo predicato Q ;
- Due regioni adiacenti devono essere diverse nel senso del predicato Q ;
- Ad esempio il predicato Q potrebbe essere il seguente:
- $Q(R_i)=\text{TRUE}$ se l'intensità media dei pixel di R_i è inferiore a m e la loro deviazione standard è minore di σ (con m e σ parametri costanti).

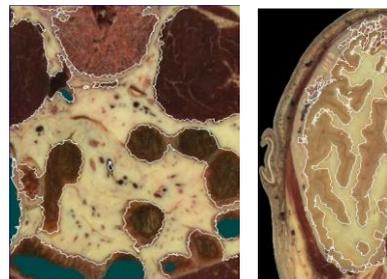
Computer Vision 2011-2012 - Prof. Sebastiano Battiato



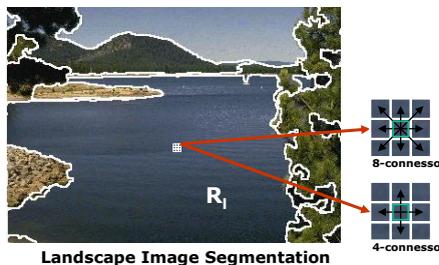
Esempi



Microarray Image Segmentation



Medical Image Segmentation



Landscape Image Segmentation

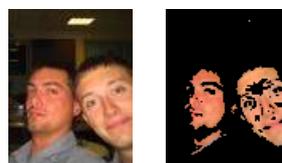
Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Esempi



Object segmentation for recognition



Skin Segmentation

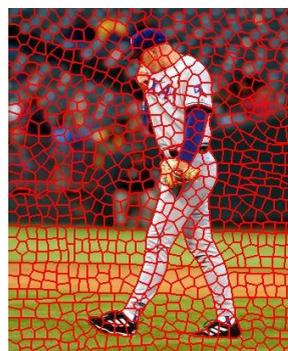
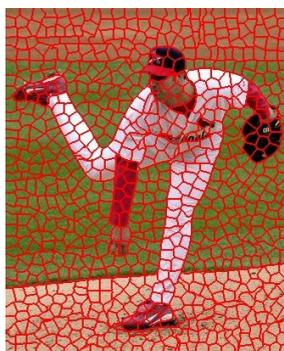
Computer Vision 2011-2012 - Prof. Sebastiano Battiato



The goals of segmentation

- Group together similar-looking pixels for efficiency of further processing
 - “Bottom-up” process
 - Unsupervised

“superpixels”



X. Ren and J. Malik. [Learning a classification model for segmentation](#)
ICCV 2003

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik

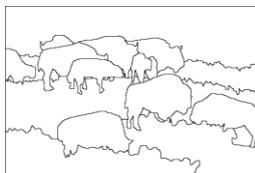


The goals of segmentation

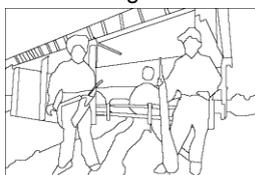
- Separate image into coherent “objects”
 - “Bottom-up” or “top-down” process?
 - Supervised or unsupervised?



image



human segmentation



Berkeley segmentation database:

<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/>

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik



Segmentation Resolution (in color space)

Si possono definire tre categorie di risoluzione:

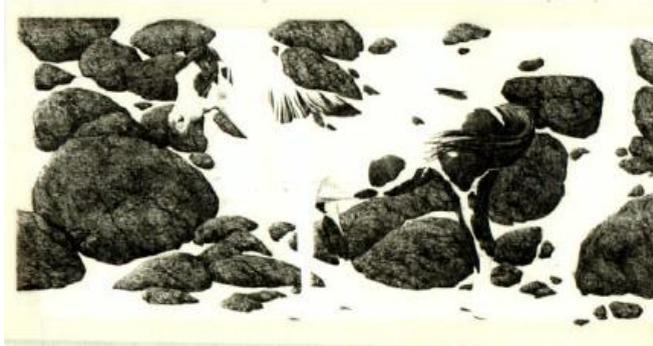
- **Under-segmentation** corrisponde alla risoluzione più bassa. L'omogeneità è definita con un grande margine di tolleranza. Soltanto i colori più significativi sono mantenuti. I contorni delle regioni sono i bordi dominanti nell'immagine.
- **Over-segmentation** corrisponde ad una risoluzione intermedia. La gamma dei colori è abbastanza ricca. L'immagine è suddivisa in piccole regioni che possono essere “fuse” attraverso controlli basati sulla conoscenza. **E' la risoluzione raccomandata quando il goal è l'object recognition.**
- **Quantization** corrisponde ad una risoluzione alta. La gamma di colori contiene tutti i colori importanti nell'immagine. Questa categoria di segmentazione è diventata importante con la diffusione delle basi di dati di immagini. Una ampia gamma di colori, possibilmente insieme alla struttura spaziale, è essenziale per le query basate sul contenuto.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Biological Segmentation in Human: Visual Perception

“To see an object in the world we must see it as something”
(L. Wittgenstein)



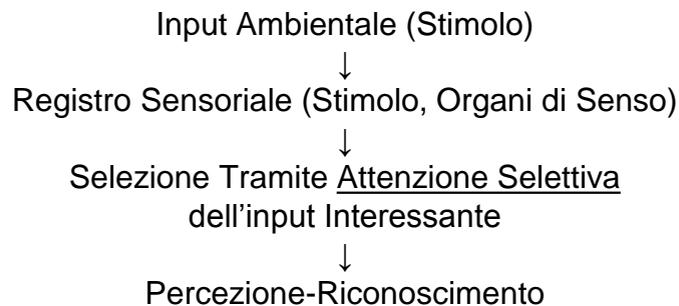
“IT is established since the Gestalt movement in psychology that perceptual grouping plays a fundamental role in human perception.” (R. Nock – F. Nielsen)

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Visual Perception

primo stadio dell'elaborazione dell'informazione che si conclude con la **Percezione:**



Human Visual Perception (Pre-Attemptive) is governed by Gestalt Principles

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Gestalt Principles of Visual Organization

Figura e sfondo

Prossimità

Chiusura

Continuazione

Similarità

Comune Destino

Parallelismo

Regione Comune

Simmetria

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Figura e Sfondo

IL RAPPORTO TRA FIGURA E SFONDO permette di “leggere” l’immagine, attraverso la separazione della figura dallo sfondo. Gli elementi *dominanti* sono percepiti come figura, il resto è percepito come sfondo.



il faro spicca maggiormente e quindi viene catalogata come figura mentre le linee orizzontali celesti sono percepite come sfondo.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Prossimità

Le colonne e le righe dominano la nostra focale. Tendiamo ad identificare come gruppi le features che sono vicine tra loro. Questa legge è da tenere in considerazione quando l'immagine gioca un ruolo importante nell'abilità di interpretare il messaggio che si nasconde dietro.

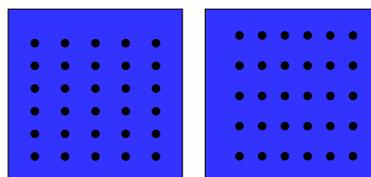


figura A

La figura A, a dimostrazione di questa legge, non viene vista come un insieme di righe ma piuttosto come un insieme di colonne.

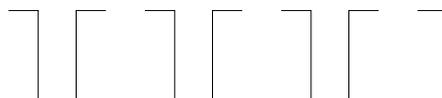


Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Chiusura

Il nostro "occhio" tende a completare gli spazi vuoti e le forme non chiuse. Noi tendiamo a vedere le immagini complete persino quando una parte dell'informazione è mancante.



In questa figura tendiamo a vedere tre rettangoli rotti ed un parentesi quadrata sulla sinistra piuttosto che tre colonne ed un parentesi quadrata sulla destra. Se eliminiamo le linee orizzontali ritorniamo alla figura che abbiamo visto nel principio di prossimità.

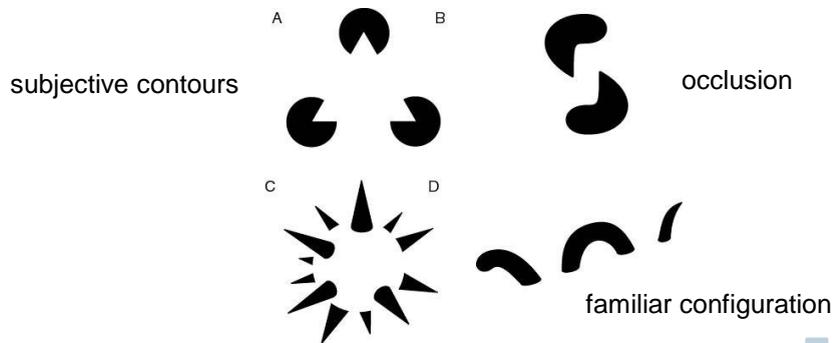


Computer Vision 2011-2012 - Prof. Sebastiano Battiato



The Gestalt school

- Elements in a collection can have properties that result from relationships
 - “The whole is greater than the sum of its parts”



http://en.wikipedia.org/wiki/Gestalt_psychology

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik



Continuazione

L'organizzazione della percezione porta lo sguardo a proseguire lungo una linea retta o una curva.



distinguiamo due linee: una da **a** a **c** e una da **b** a **d**.

In realtà questo grafico potrebbe rappresentare un altro insieme di linee: da **a** a **d** e da **b** a **c**. Tuttavia, è più probabile che si tenda ad identificare il primo gruppo di queste linee che hanno una *migliore continuazione* rispetto al secondo dove è presente un'ovvia torsione.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

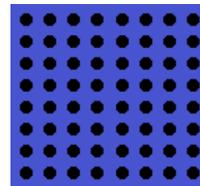
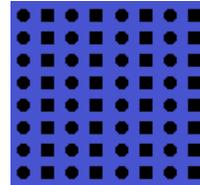


Similarità

Elementi visivi simili vengono raggruppati sulla base della forma, della grandezza, del colore o della direzione.

I cerchi e i quadrati sono raggruppati in maniera naturale; percepiamo colonne alternate di cerchi e quadrati.

Se non fossero presenti le due caratteristiche ricorrenti percepiremmo la figura come righe o colonne



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Parallelismo

Linee parallele tendono ad essere raggruppate insieme

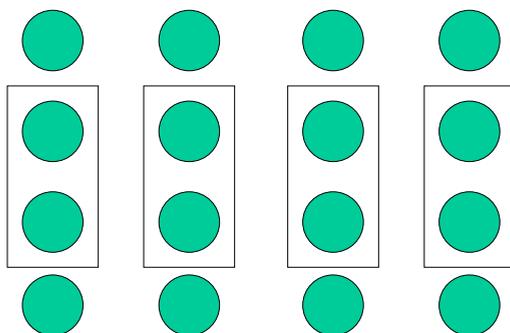


Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Regione Comune

Figure posizionate all'interno della stessa regione chiusa tendono ad essere percepite insieme



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Simmetria

Percepire l'intero di una figura rispetto alle parti singole che la costituiscono



Osservando la figura si vedono due rombi sovrapposti o tre oggetti: un piccolo rombo e due oggetti irregolari sopra e sotto di esso.

Secondo il principio di simmetria, si percepiranno probabilmente due rombi sovrapposti

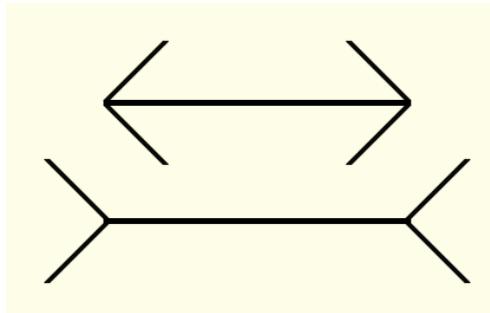
Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Inspiration from psychology

- The Gestalt school: Grouping is key to visual perception

The Muller-Lyer illusion



http://en.wikipedia.org/wiki/Gestalt_psychology

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Emergence



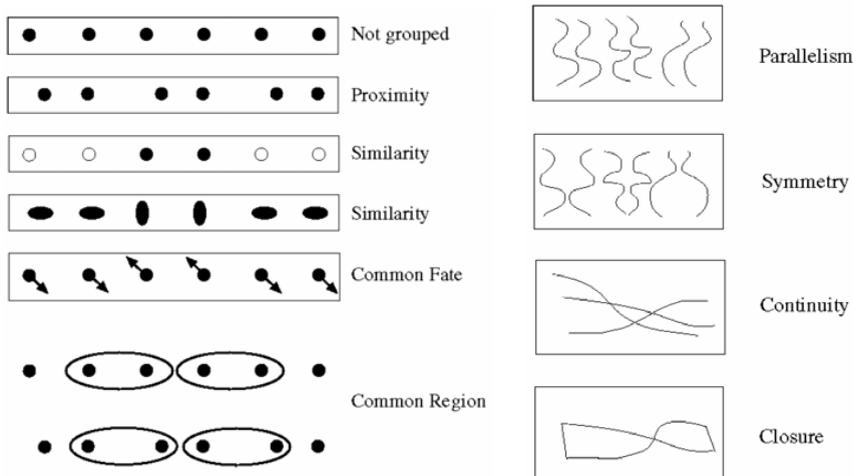
http://en.wikipedia.org/wiki/Gestalt_psychology

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik



Gestalt factors



- These factors make intuitive sense, but are very difficult to translate into algorithms

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik



Gestalt Principles e Strategie di Segmentazione

Molte strategie di segmentazione sono riconducibili ai principi Gestalt

Gestalt Principles	Strategie di Segmentazione
Figure/Sfondo	Tresholding
Prossimità, Similarità	Clustering, Region Growing, Region Merging
Continuità e Chiusura	Morphological Methods
Similarità	Statistical Methods
Continuità e chiusura	Edge-Detection Based

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di segmentazione

Nel corso degli anni sono state sviluppate svariate tecniche di segmentazione. Tuttavia una soluzione generale al problema non è stata trovata. Alcune delle principali strategie di segmentazione sono:

Edge-based;
Thresholding;
Region Merging;
Region Splitting and Merging.

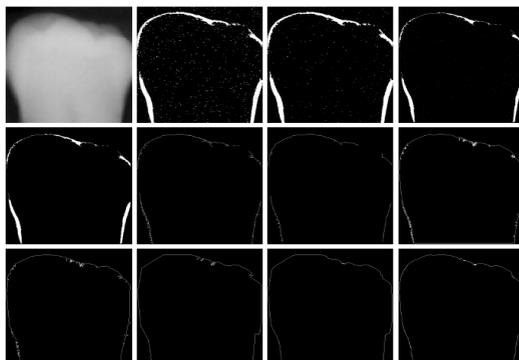
...

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di segmentazione: edge-based

Estrazione degli edge (Sobel, Canny, ...)
Edge Linking e Boundary detection (trasformata di Hough,...).



Esempio di immagine segmentata con un approccio edge-based.

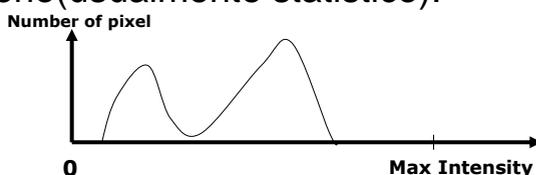
Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di Segmentazione: Thresholding

Ogni pixel di una immagine a toni di grigio è caratterizzato da un valore di luminanza. L'algoritmo, di base, fissa una soglia di luminanza, allo scopo di distinguere tra due regioni(Sfondo/Oggetti).

La soglia può essere fissata arbitrariamente, o in maniera automatica in base ad un certo criterio(usualmente statistico).



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di segmentazione: thresholding (1)

I metodi di segmentazione basati sull'analisi dell'istogramma sono spesso utilizzati grazie alla loro semplicità implementativa ed efficienza computazionale.

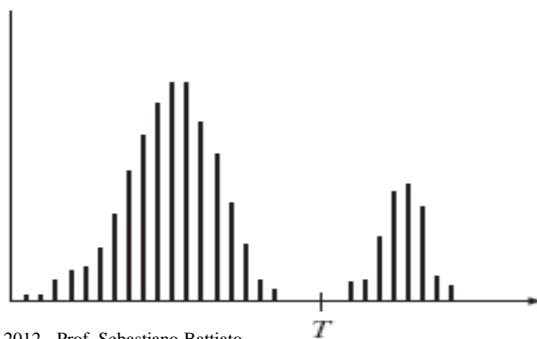
Queste tecniche calcolano un istogramma a partire dai pixel (es. intensità) e utilizzano i sui picchi e le sue valli per localizzare i cluster dell'immagine.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di segmentazione: thresholding (2)

Supponiamo di avere un oggetto chiaro su sfondo scuro e che il suo istogramma sia quello mostrato in figura. I pixel dell'oggetto e del background sono raggruppati in due mode dominanti.



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di segmentazione: thresholding (3)

Scelta una soglia T che separa le due mode, un punto (x,y) tale che $f(x,y) > T$ sarà un punto dell'oggetto, altrimenti verrà assegnato allo sfondo.

Se T è una costante che può essere applicata all'intera immagine, si parla di sogliatura globale. Se il valore di T varia sull'immagine si utilizza il termine sogliatura variabile.

Nel caso in cui sia necessario discriminare più di due classi la segmentazione è piuttosto complessa. In tali casi, tipicamente si ottengono migliori risultati con altri approcci (soglia variabile e region merging).

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Strategie di segmentazione: thresholding (4)

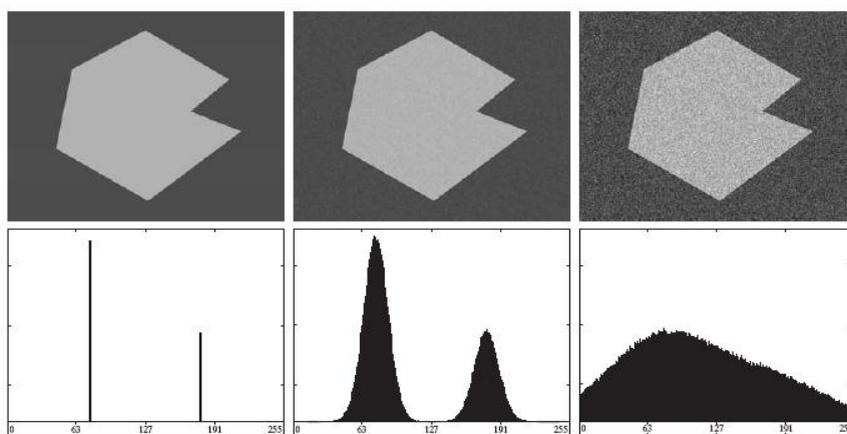
La buona riuscita degli algoritmi basati sugli istogrammi dipende dalla larghezza e dalla profondità delle valli che separano le mode dell'istogramma. I fattori che influenzano le proprietà delle valli sono:

- La separazione tra i picchi;
- Il rumore presente nell'immagine;
- La dimensione relativa dell'oggetto rispetto allo sfondo;
- L'uniformità della sorgente luminosa;
- L'uniformità delle proprietà di riflettanza dell'immagine.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Ruolo del rumore

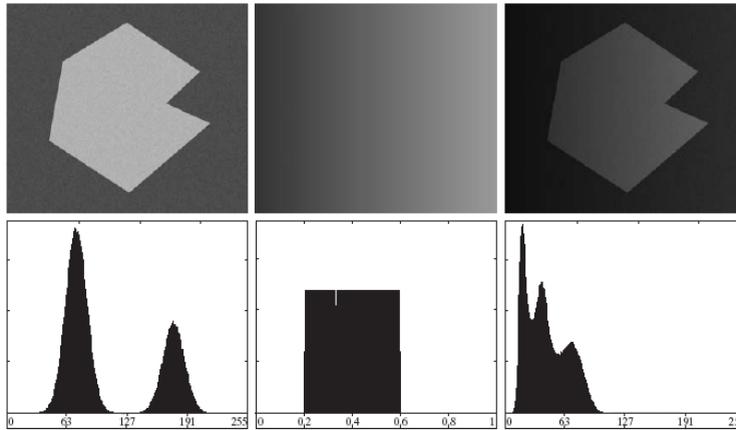


All'aumentare del rumore presente nell'immagine, il processo di segmentazione diventa sempre più complesso.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Ruolo dell'illuminazione



Una sorgente luminosa non uniforme può rendere molto complessa la segmentazione. In figura viene mostrato un esempio di immagine illuminata con una sorgente non uniforme (una rampa) ed il suo istogramma.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Threshold Selection: Peak and Valley method

Histogramma (h): distribuzione di frequenza dei livelli di grigio presenti nell'immagine $I_{rc}(i,j)$

$h(g)$ = numero di pixel in I_{rc} il cui livello di grigio è g

1. Trova i due picchi più prominenti di h
 - g è un picco se $h(g) > h(g \pm Dg)$, $Dg = 1, \dots, k$

Siano g_1 e g_2 i due picchi più alti con $g_1 < g_2$

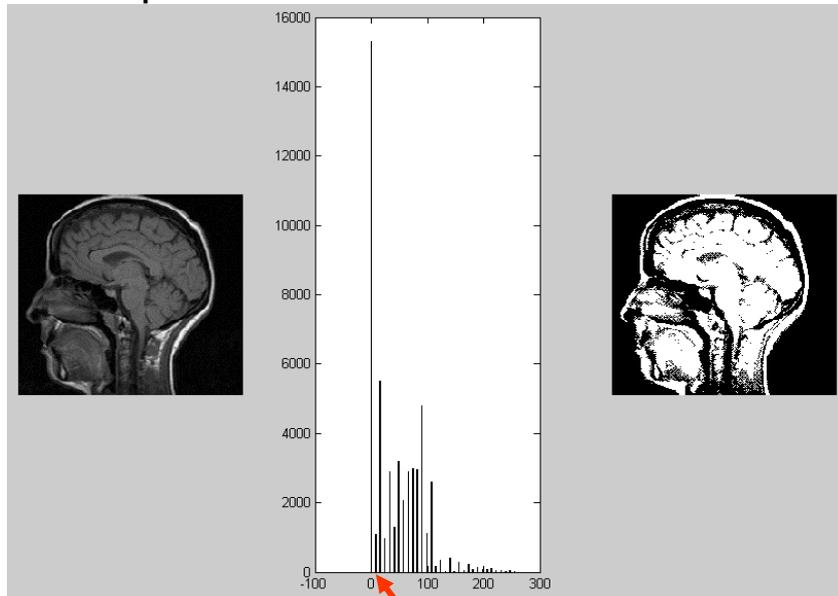
2. Trova la valle più profonda, g , tra g_1 e g_2
 - g è una valle se $h(g) \leq h(g')$ per ogni g' , con $g, g' \in [g_1, g_2]$

Usa g come threshold

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Esempio



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

IMAGE PROCESSING LABORATORY

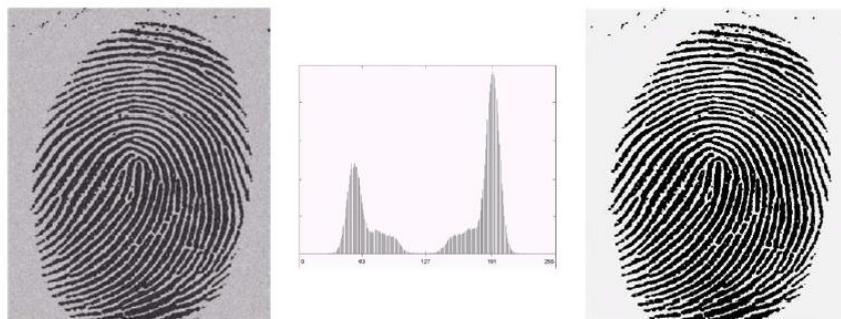
Global threshold selection

1. Seleziona una soglia iniziale T (es: media dei grigi presenti)
2. Segmenta l'immagine usando T
 - Questo passo produce due gruppi di pixel
 - G_1 : tutti i pixel hanno valore minore a T
 - G_2 : tutti i pixel hanno valore maggiore uguale a T
3. Computa la media sui valori di grigio dei pixel presenti in G_1 e G_2 (μ_1, μ_2)
4. Computa la nuova soglia threshold
 - $T = \frac{1}{2} (\mu_1 + \mu_2)$
5. Ripeti gli step da 2-4 fino a convergenza
 - Es. Convergenza: T non cambia in due ripetizioni successive

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

IMAGE PROCESSING LABORATORY

Esempio



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Metodo di Otsu

Tecnica automatica di selezione della soglia T .
Il metodo massimizza la varianza inter-classe;
E' basato solo su operazioni effettuate
sull'istogramma dell'immagine (vettore 1-D).

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Metodo di Otsu (2)

Consideriamo un'immagine di dimensioni $M \times N$ con L livelli distinti di intensità e sia n_i il numero di pixel di intensità i . Il suo istogramma normalizzato ha componenti $p_i = n_i / MN$.

Supponiamo di selezionare una soglia $T(k) = k$, $0 < k < L-1$, e di dividere in base ad essa l'immagine in due classi, C_1 (tutti i pixel con intensità $[0, k]$) e C_2 (tutti i pixel con intensità $[k+1, L-1]$).

Con tale soglia, la probabilità che un pixel sia assegnato alla classe C_1 è:

$$P_1(k) = \sum_{i=0}^k p_i$$

Per la classe C_2 si ha:

$$P_2(k) = \sum_{i=k+1}^{L-1} p_i = 1 - P_1(k)$$

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Metodo di Otsu (3)

Il valore medio di intensità dei pixel appartenenti alla classe C_1 è:

$$m_1(k) = \sum_{i=0}^k iP(i / C_1) = \sum_{i=0}^k iP \frac{P(C_1 / i)P(i)}{P(C_1)} = \frac{1}{P_1(k)} \sum_{i=0}^k ip_i$$

In maniera simile si

ricava:

$$m_2(k) = \sum_{i=k+1}^{L-1} iP(i / C_2) = \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} ip_i$$

Si definisce media cumulativa fino al livello k :

$$m(k) = \sum_{i=0}^k ip_i$$

La media delle intensità dell'intera immagine è data da:

$$m_G = \sum_{i=0}^{L-1} ip_i$$

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Metodo di Otsu (4)

Per valutare la bontà della soglia k si utilizza la metrica normalizzata:

$$\eta = \frac{\sigma_B^2}{\sigma_G^2}$$

Con σ_G^2 varianza globale dei pixel dell'immagine:

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i$$

e σ_B^2 varianza inter-classe:

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2$$

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Metodo di Otsu (5)

La varianza inter-classe può essere riscritta come:

$$\sigma_B^2 = P_1 P_2 (m_1 - m_2)^2 = \frac{(m_G P_1 - m)^2}{P_1(1 - P_1)}$$

Più distanti sono le due medie, più la varianza inter-classe è elevata. Questa formulazione della varianza inter-classe permette una più efficiente implementazione.

La soglia ottimale viene dunque calcolata:

$$\sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k)$$

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



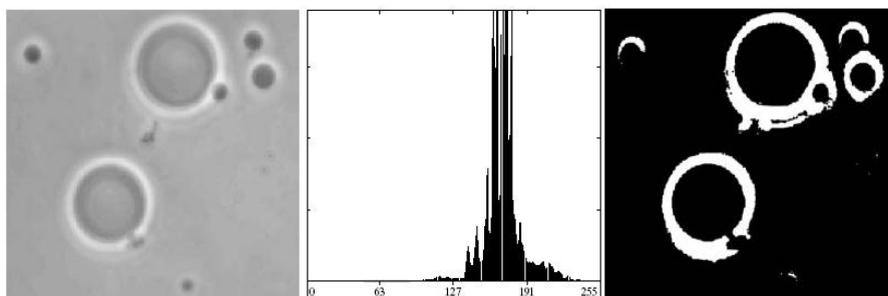
Metodo di Otsu (6)

- L'algoritmo di Otsu può essere riassunto come segue:
- Calcolare l'istogramma normalizzato dell'immagine;
 - Calcolare le somme cumulative $P_1(k)$ per $k=0, 1, \dots, L-1$;
 - Calcolare le medie cumulative $m(k)$ per $k=0, 1, \dots, L-1$;
 - Calcolare la media globale delle intensità, m_G ;
 - Calcolare la varianza inter-classe per $k=0, 1, \dots, L-1$;
 - Ottenere la soglia k^* che massimizza la varianza inter-classe; Se il massimo non è unico, ricavare k^* come media dei valori di k corrispondenti ai vari massimi trovati;
 - Ricavare la misura di separabilità η^* per $k=k^*$.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Metodo di Otsu (esempio)



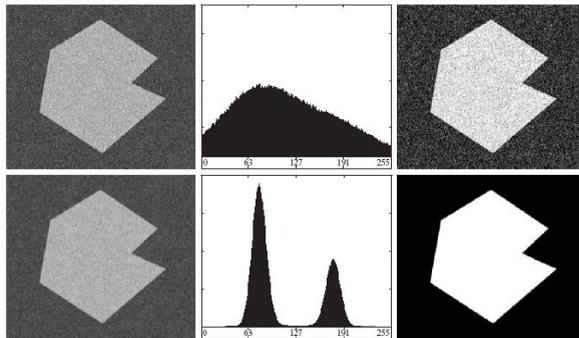
Esempio di immagine segmentata con il metodo di Otsu.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Ruolo del rumore

In alcuni casi l'immagine presenta un livello di rumore tale da rendere complessa la segmentazione tramite sogliatura. Spesso l'applicazione di un filtro di smoothing permette di ridurre il problema.



L'applicazione di una maschera di smoothing ha semplificato la segmentazione rendendo l'istogramma bimodale



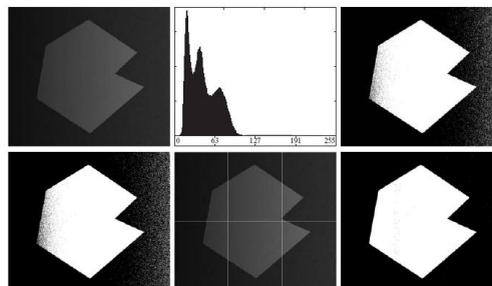
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Illuminazione non uniforme

Nel caso in cui l'immagine sia illuminata in maniera non uniforme o abbia delle non uniformità nella riflettanza, la segmentazione tramite thresholding può risultare piuttosto complessa.

Una soluzione semplice al problema consiste nel partizionare l'immagine in rettangoli non sovrapposti e su di essi effettuare la segmentazione.

L'istogramma relativo ad ogni rettangolo è bimodale, e permette dunque la segmentazione dell'immagine.



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Sogliatura variabile (cenni)

Quando il valore della soglia T varia sull'immagine si utilizza il termine sogliatura variabile. Alcune tecniche sono le seguenti:

Partizionamento dell'immagine: l'immagine viene divisa in rettangoli non sovrapposti e su di essi vengono calcolate le soglie per la segmentazione;

Sogliature basate su proprietà locali dell'immagine: viene calcolata una soglia per ogni punto in base a delle proprietà calcolate in un suo intorno (es. media, varianza...);

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Region Merging

A partire da un pixel detto *seed* si agglomerano ad esso i pixel a lui vicini che soddisfano un certo criterio di omogeneità formando così una regione.

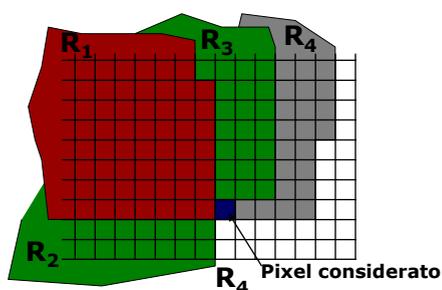
Combinando successivi processi di *growing*, o procedendo con *growing* simultaneo da più *seed*, si ottiene la segmentazione dell'intera immagine.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Region Growing

Ad ogni istante viene preso in considerazione un pixel che non è stato ancora allocato ma che è adiacente ad almeno una regione; il pixel è allocato alla regione adiacente che risulta più simile secondo il criterio scelto.



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Region Growing Algorithm

1. Scegli un seed (pixel) iniziale
2. Seleziona i pixel vicini (connessi) e fai il merge se la condizione di omogeneità scelta è soddisfatta
3. Se la regione non cresce, seleziona un altro seed e ripeti dal punto 2 finché tutti i pixel non sono stati allocati ad una regione, altrimenti vai al punto 2
4. Rimuovi le regioni molto piccole (passo opzionale)

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Seeded Region Growing Algorithm

Viene dato in input il numero di *seeds* (pixel di partenza utilizzati per far crescere le regioni).

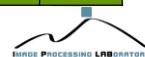
L'algoritmo quindi procede autonomamente facendo crescere simultaneamente le regioni, finché tutti i pixel nell'immagine sono stati racchiusi in una regione. Per ogni passo tutti i pixel che non sono stati ancora allocati, ma che hanno almeno un vicino allocato, vengono presi in considerazione: tra tutte le regioni confinanti al pixel considerato, l'algoritmo seleziona quella i cui pixel hanno in media la minore differenza (es: in termini di livelli di grigio) rispetto al pixel preso in considerazione.



Esempio

115	115	115	129	129	115	115	115	129	129
115	115	115	129	129	115	115	115	129	129
117	117	117	10	10	117	117	117	10	10
117	117	117	10	10	117	117	117	10	10
117	117	117	10	10	117	117	117	10	10

115	115	115	129	129	115	115	115	129	129
115	115	115	129	129	115	115	115	129	129
117	117	117	10	10	117	117	117	10	10
117	117	117	10	10	117	117	117	10	10
117	117	117	10	10	117	117	117	10	10



Region Splitting and Merging (1)

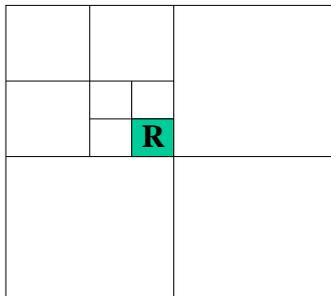
E' possibile eseguire una segmentazione partizionando (splitting) ricorsivamente una immagine, fino ad ottenere componenti uniformi.

Si dovrà effettuare una successiva operazione di aggregazione (merging) delle regioni adiacenti che dovessero risultare compatibili in base ad un criterio di fusione.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

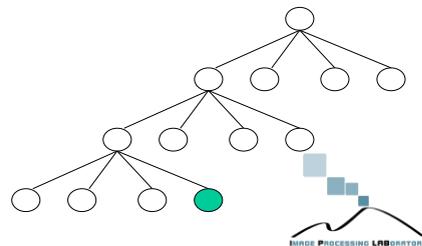


Region Splitting and Merging (2)



La suddivisione ricorsiva dell'immagine in quadranti viene rappresentata con una struttura ad albero chiamato quad tree: ogni nodo contiene le informazioni relative a ciascun quadrante e i suoi figli sono associati ai quadranti in cui è ulteriormente suddiviso. Un nodo foglia è un quadrante sufficientemente uniforme da non richiedere ulteriori partizionamenti

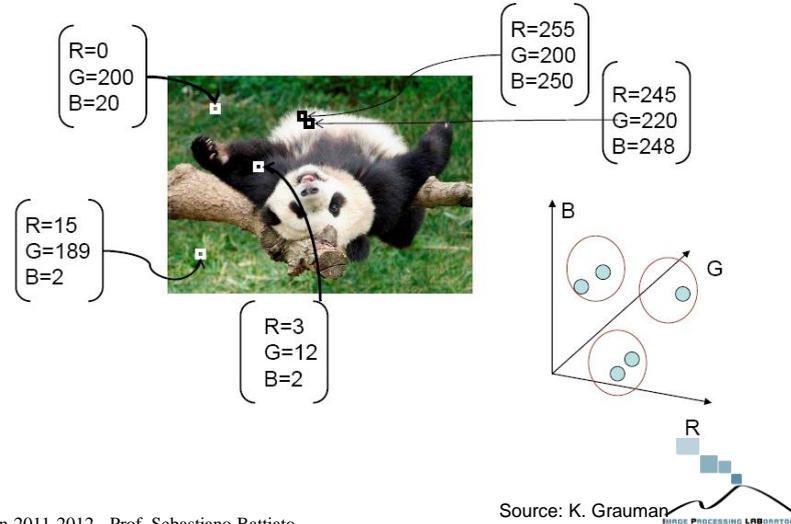
Dopo la fase di splitting si procederà alla fase di merging delle regioni adiacenti "compatibili"; regioni adiacenti verranno aggregate in una unica regione se quest'ultima risulterà sufficientemente uniforme.



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Segmentation as clustering

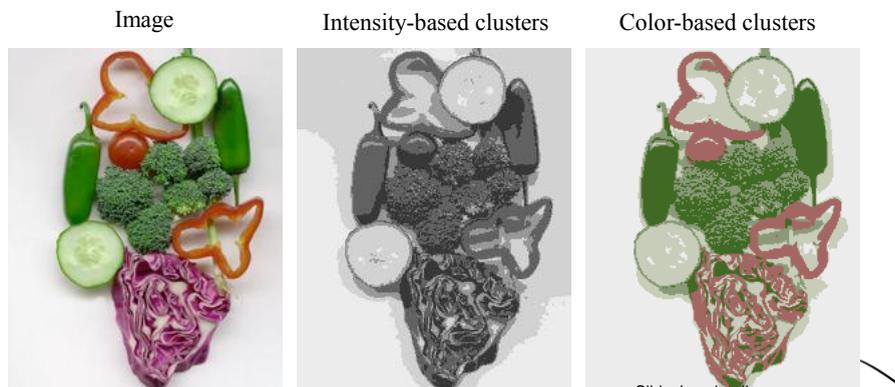
- Cluster similar pixels (features) together



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Segmentation as clustering

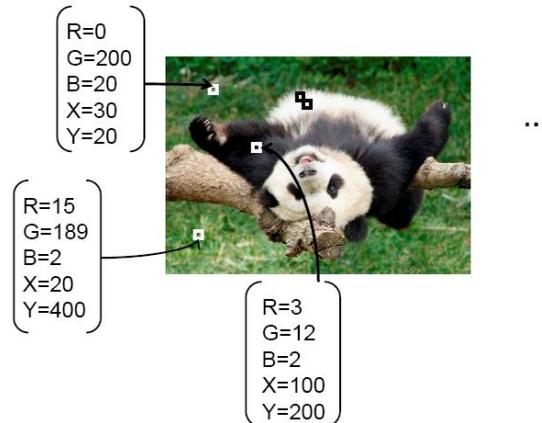
- K-means clustering based on intensity or color is essentially vector quantization of the image attributes
 - Clusters don't have to be spatially coherent



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Segmentation as clustering

- Cluster similar pixels (features) together

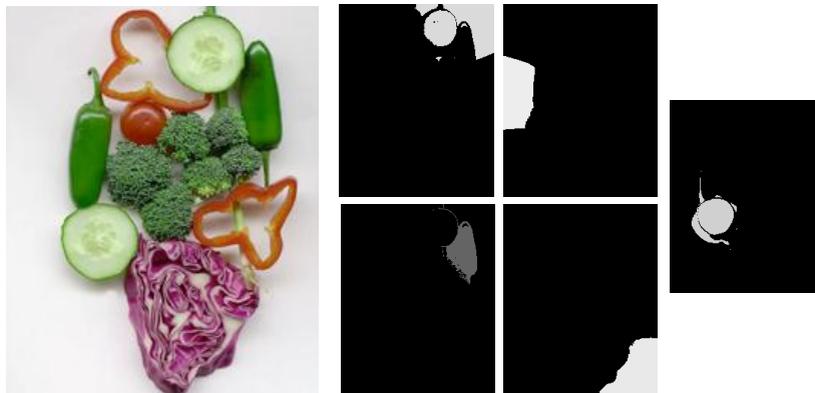


Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Segmentation as clustering

Clustering based on (r,g,b,x,y) values enforces more spatial coherence

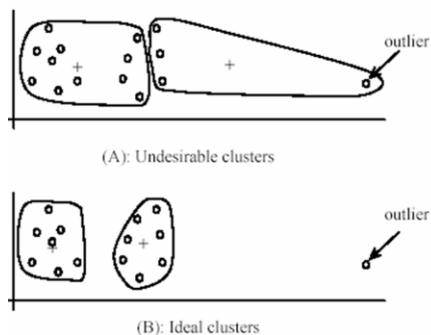


Computer Vision 2011-2012 - Prof. Sebastiano Battiato



K-Means for segmentation

- Pros
 - Very simple method
 - Converges to a local minimum of the error function
- Cons
 - Memory-intensive
 - Need to pick K
 - Sensitive to initialization
 - Sensitive to outliers
 - Only finds “spherical” clusters

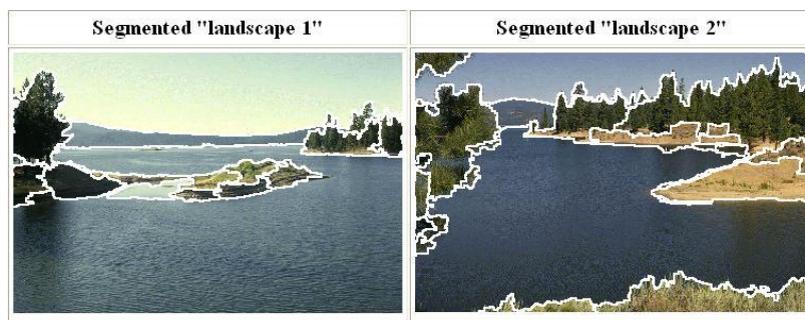


Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik 

Mean shift clustering and segmentation

- An advanced and versatile technique for clustering-based segmentation



<http://www.caip.rutgers.edu/~comanici/MSPAMI/msPamiResults.html>

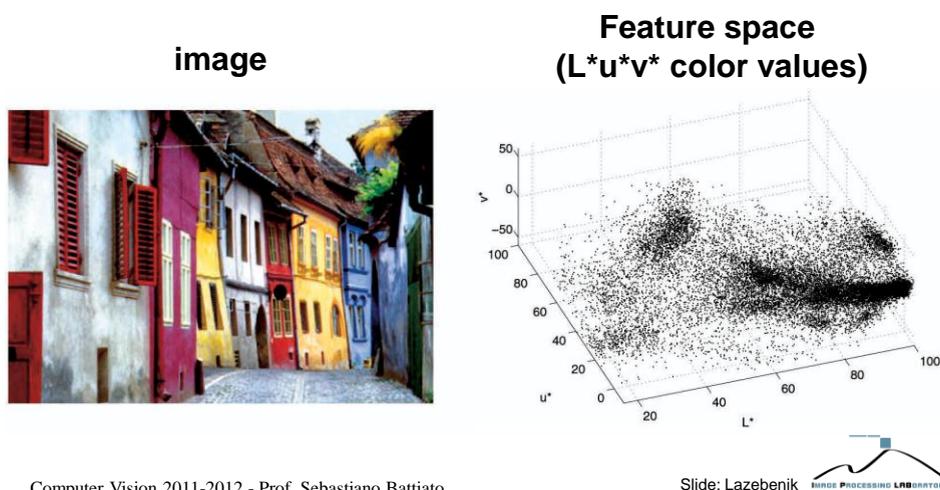
D. Comaniciu and P. Meer, [Mean Shift: A Robust Approach toward Feature Space Analysis](#), PAMI 2002.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

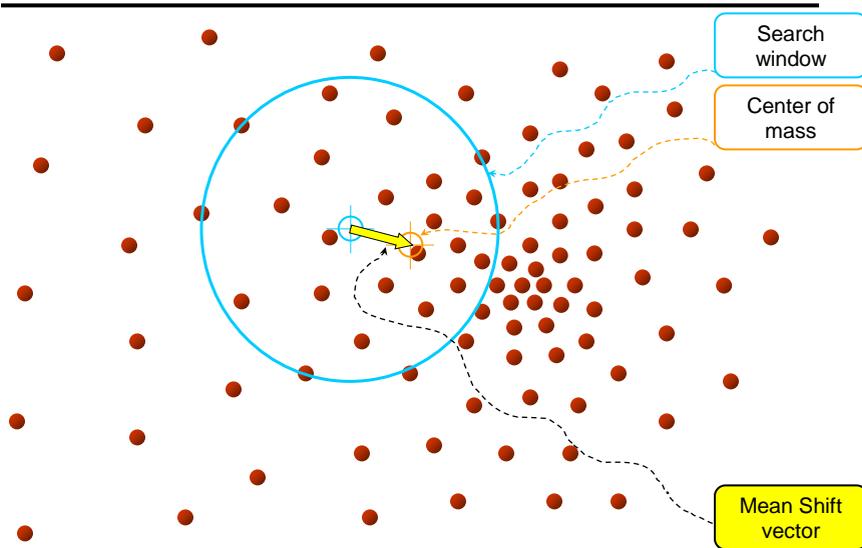
Slide: Lazebenik 

Mean shift algorithm

- The mean shift algorithm seeks *modes* or local maxima of density in the feature space



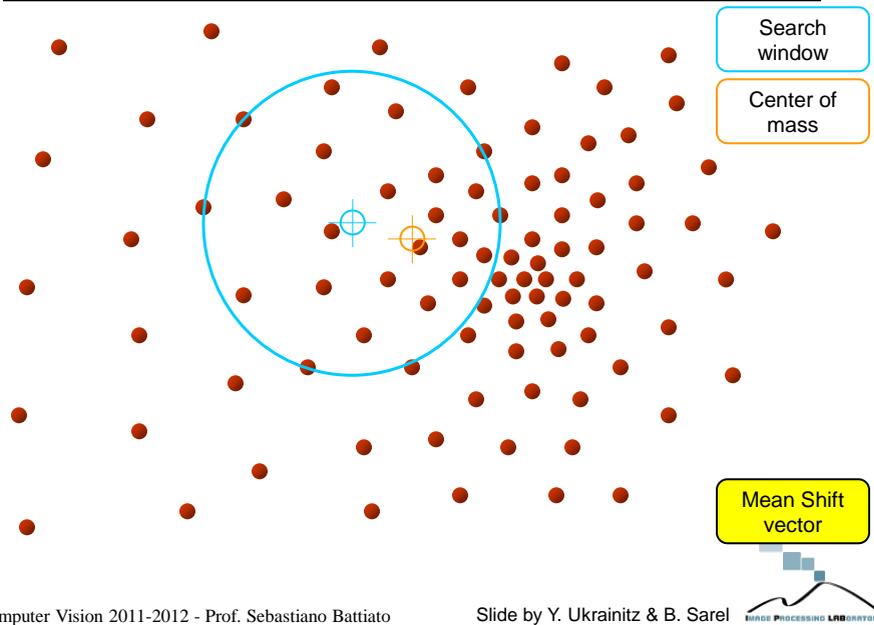
Mean shift



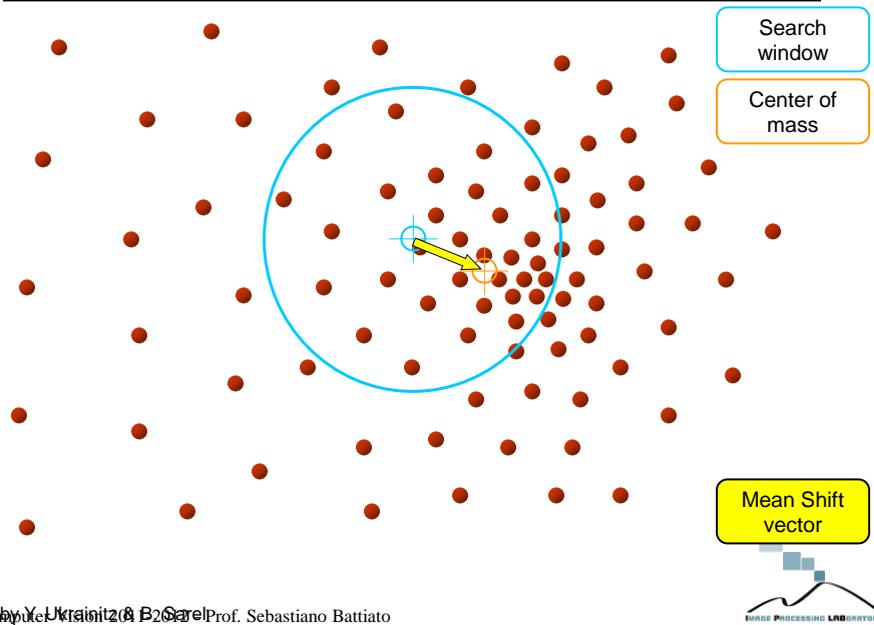
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide by Y. Ukrainitz & B. Sarel 

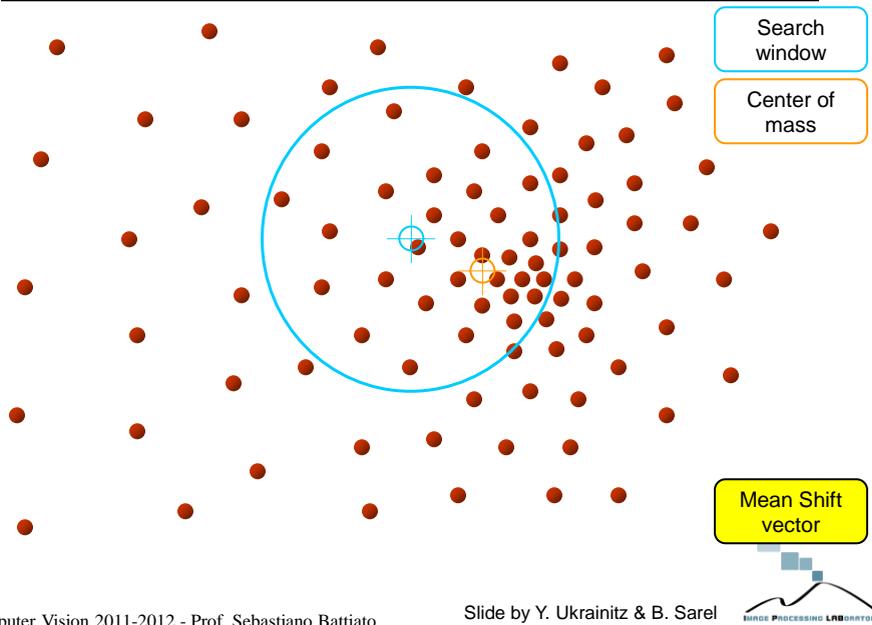
Mean shift



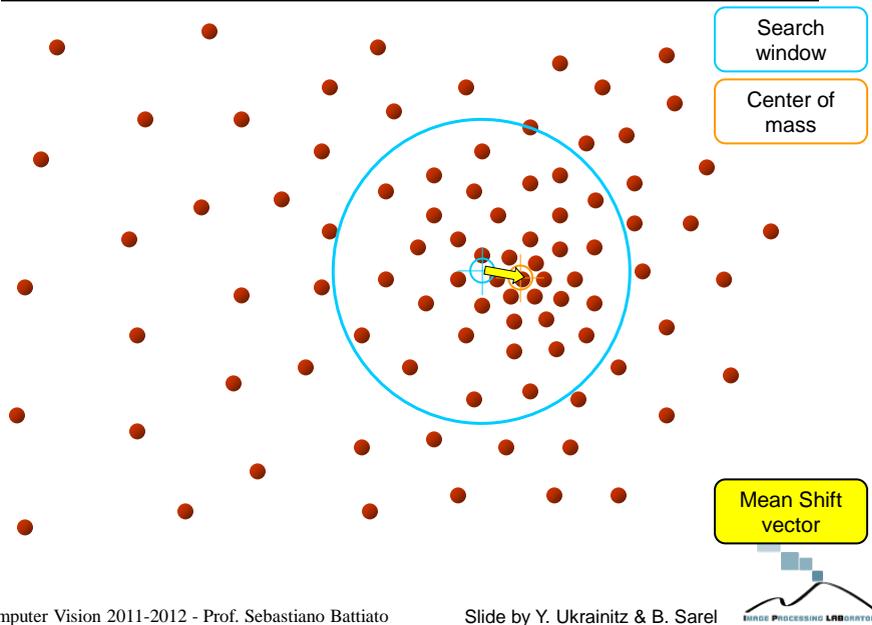
Mean shift



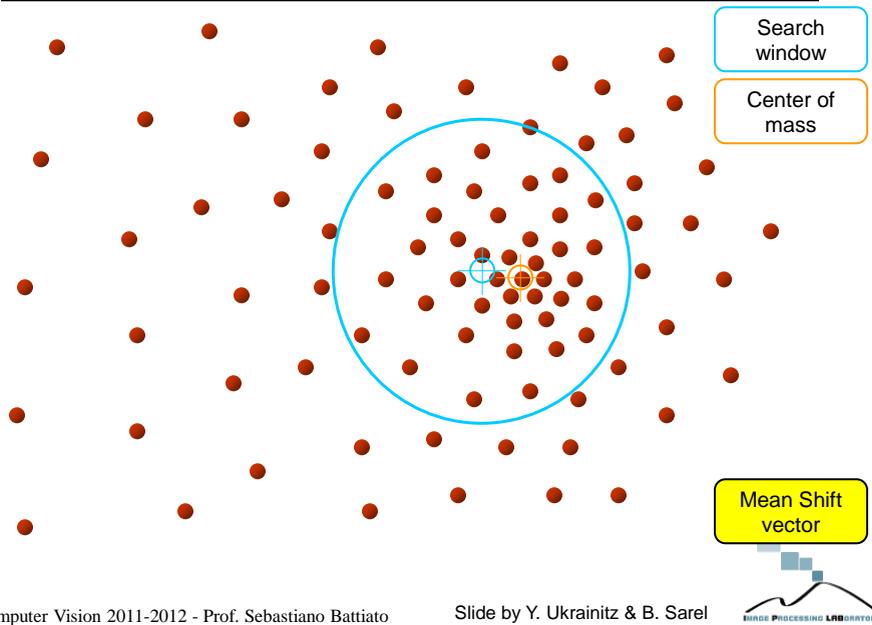
Mean shift



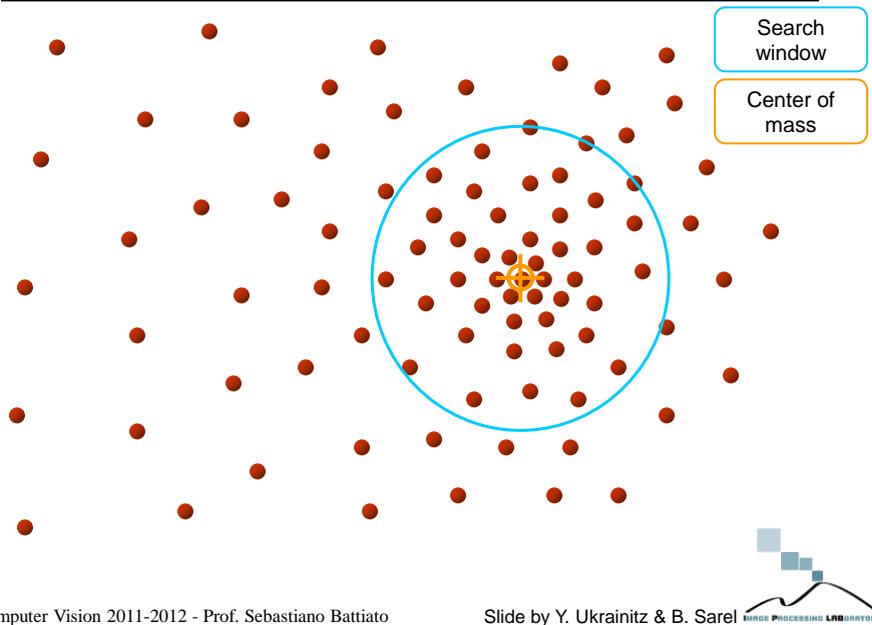
Mean shift



Mean shift

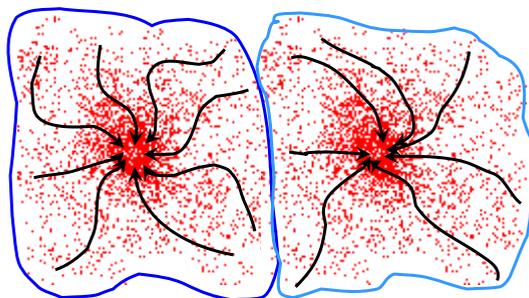


Mean shift



Mean shift clustering

- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region for which all trajectories lead to the same mode



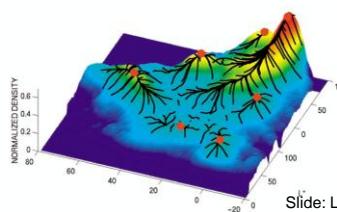
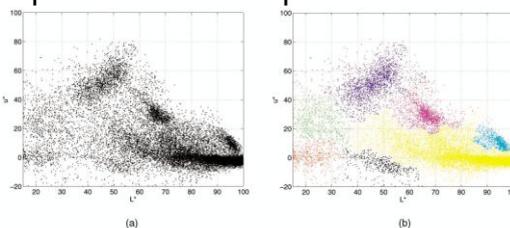
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide by Y. Ukrainitz & B. Saref



Mean shift clustering/segmentation

- Find features (color, gradients, texture, etc)
- Initialize windows at individual feature points
- Perform mean shift for each window until convergence
- Merge windows that end up near the same “peak” or mode



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik

Mean shift segmentation results



<http://www.caip.rutgers.edu/~comanici/MSPAMI/msPamiResults.html>
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik 

More results



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik 

More results



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik



Mean shift pros and cons

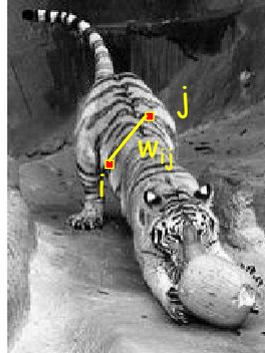
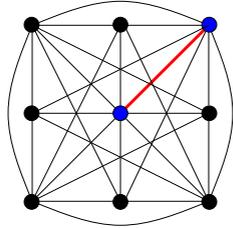
- Pros
 - Does not assume spherical clusters
 - Just a single parameter (window size)
 - Finds variable number of modes
 - Robust to outliers
- Cons
 - Output depends on window size
 - Computationally expensive
 - Does not scale well with dimension of feature space

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik



Images as graphs

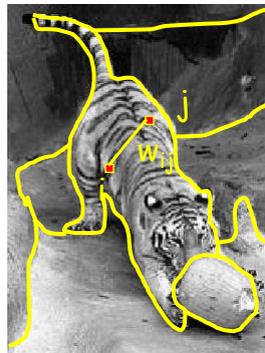
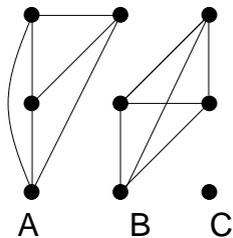


- Node for every pixel
- Edge between every pair of pixels (or every pair of “sufficiently close” pixels)
- Each edge is weighted by the *affinity* or similarity of the two nodes

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Source: S. Seitz 

Segmentation by graph partitioning



- Break Graph into Segments
 - Delete links that cross between segments
 - Easiest to break links that have low affinity
 - similar pixels should be in the same segments
 - dissimilar pixels should be in different segments

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Source: S. Seitz 

Measuring affinity

- Suppose we represent each pixel by a feature vector \mathbf{x} , and define a distance function appropriate for this feature representation
- Then we can convert the distance between two feature vectors into an affinity with the help of a generalized Gaussian kernel:

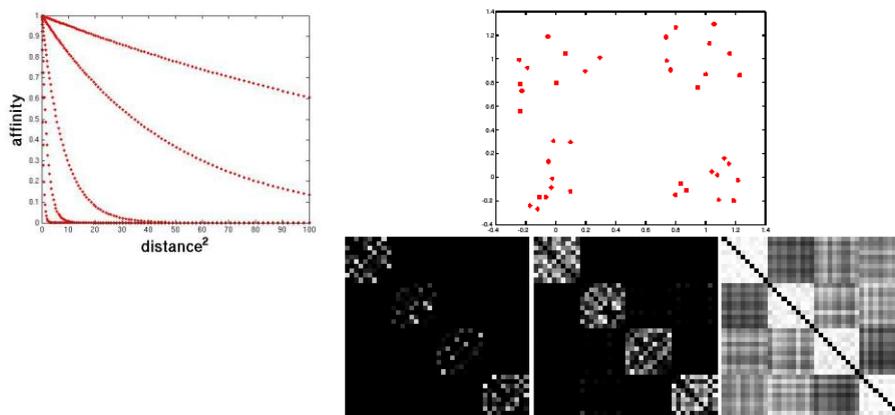
$$\exp\left(-\frac{1}{2\sigma^2} \text{dist}(\mathbf{x}_i, \mathbf{x}_j)^2\right)$$

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Scale affects affinity

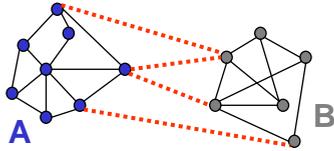
- Small σ : group only nearby points
- Large σ : group far-away points



Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Graph cut



- Set of edges whose removal makes a graph disconnected
- Cost of a cut: sum of weights of cut edges
- A graph cut gives us a segmentation
 - What is a “good” graph cut and how do we find one?

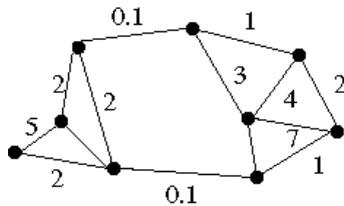
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Source: S. Seitz 

Minimum cut

- We can do segmentation by finding the *minimum cut* in a graph
 - Efficient algorithms exist for doing this

Minimum cut example



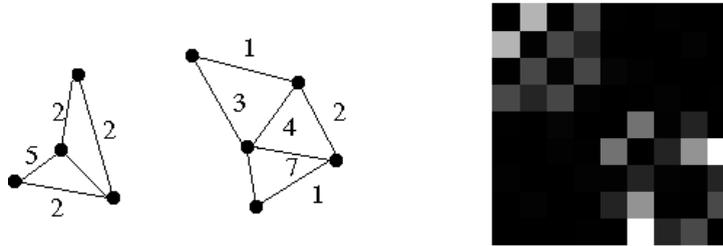
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik 

Minimum cut

- We can do segmentation by finding the *minimum cut* in a graph
 - Efficient algorithms exist for doing this

Minimum cut example

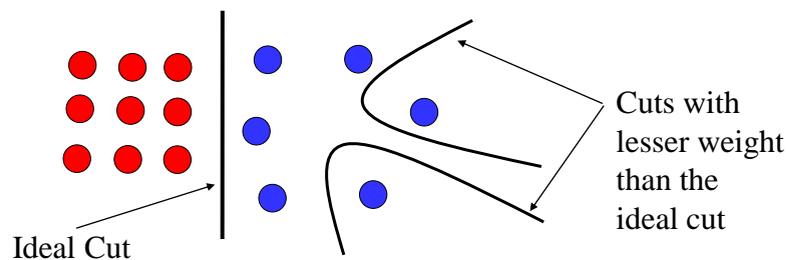


Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik 

Normalized cut

- Drawback: minimum cut tends to cut off very small, isolated components



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

* Slide from Khurram Hassan-Shafique CAP5415  Computer Vision 2003

Normalized cut

- Drawback: minimum cut tends to cut off very small, isolated components
- This can be fixed by normalizing the cut by the weight of all the edges incident to the segment
- The *normalized cut* cost is:

$$\frac{w(A, B)}{w(A, V)} + \frac{w(A, B)}{w(B, V)}$$

$w(A, B)$ = sum of weights of all edges between A and B

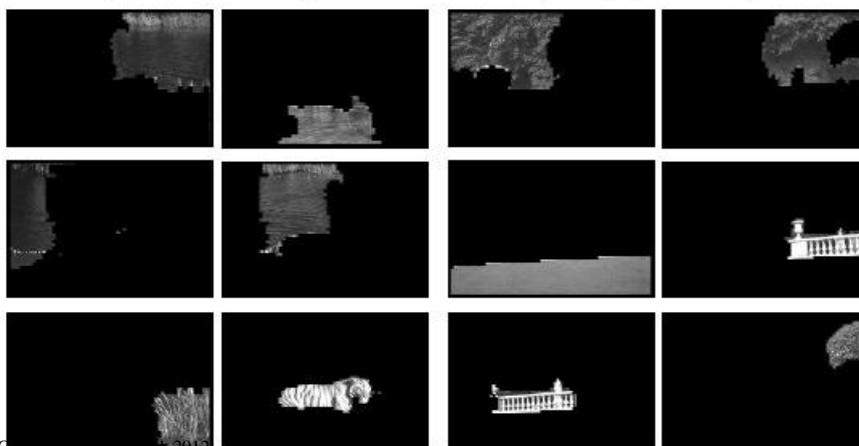
J. Shi and J. Malik. [Normalized cuts and image segmentation](#). PAMI 2000

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebernik



Example result



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

ITV

Challenge

- How to segment images that are a “mosaic of textures”?

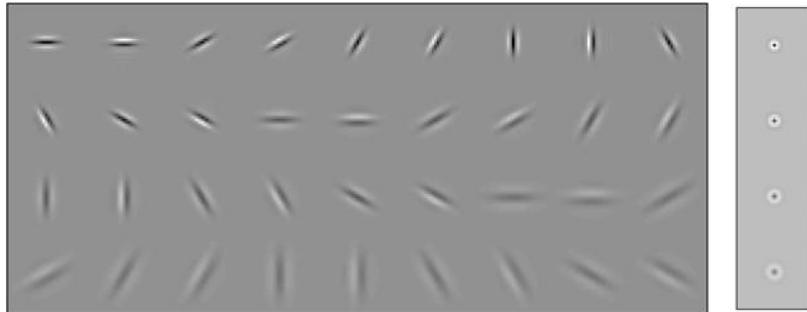


Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik 

Using texture features for segmentation

- Convolve image with a bank of filters



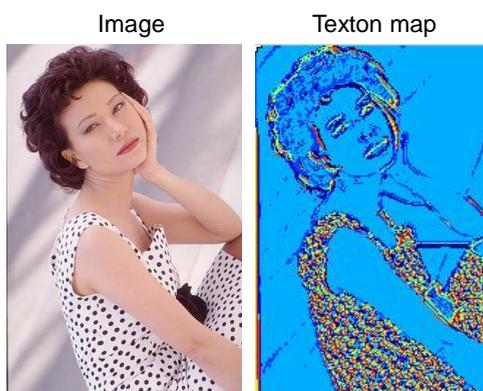
J. Malik, S. Belongie, T. Leung and J. Shi. "[Contour and Texture Analysis for Image Segmentation](#)". IJCV 43(1),7-27,2001.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik 

Using texture features for segmentation

- Convolve image with a bank of filters
- Find *textons* by clustering vectors of filter bank outputs



J. Malik, S. Belongie, T. Leung and J. Shi. "[Contour and Texture Analysis for Image Segmentation](#)". IJCV 43(1),7-27,2001.

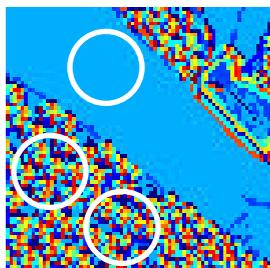
Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik



Using texture features for segmentation

- Convolve image with a bank of filters
- Find *textons* by clustering vectors of filter bank outputs
- The final texture feature is a texton histogram computed over image windows at some "local scale"



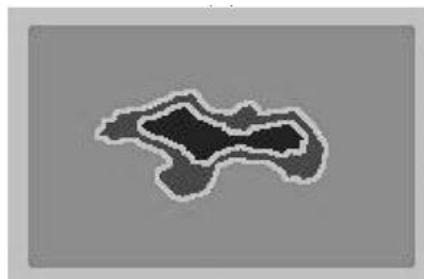
J. Malik, S. Belongie, T. Leung and J. Shi. "[Contour and Texture Analysis for Image Segmentation](#)". IJCV 43(1),7-27,2001.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik



Pitfall of texture features



- Possible solution: check for “intervening contours” when computing connection weights

J. Malik, S. Belongie, T. Leung and J. Shi. "[Contour and Texture Analysis for Image Segmentation](#)". IJCV 43(1),7-27,2001.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebenik

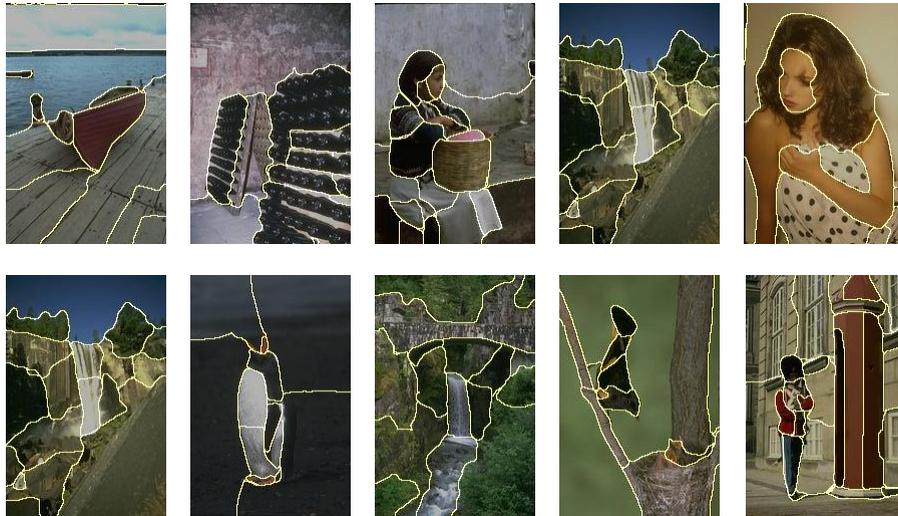


Example results



Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Results: Berkeley Segmentation Engine



<http://www.cs.berkeley.edu/~fowlkes/BSE/>

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



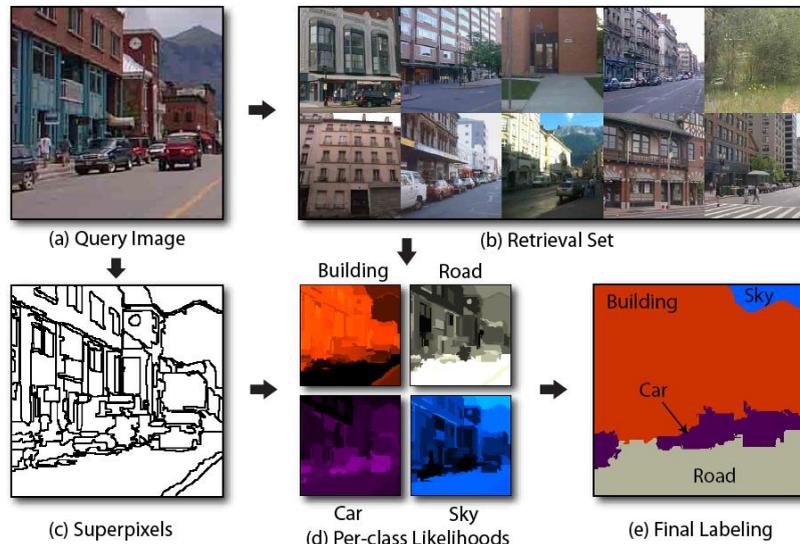
Normalized cuts: Pro and con

- Pros
 - Generic framework, can be used with many different features and affinity formulations
- Cons
 - High storage requirement and time complexity
 - Bias towards partitioning into equal segments

Computer Vision 2011-2012 - Prof. Sebastiano Battiato



Segments as primitives for recognition

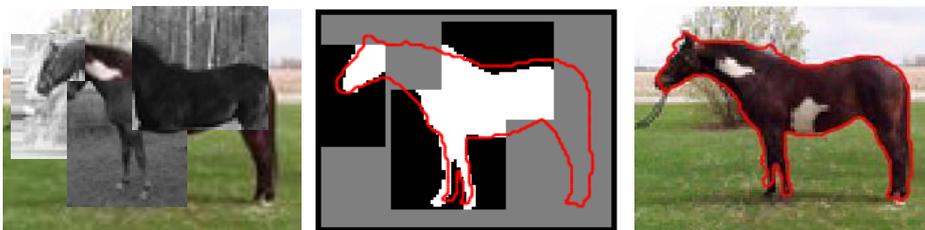


J. Tighe and S. Lazebnik, ECCV 2010

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebnik 

Top-down segmentation



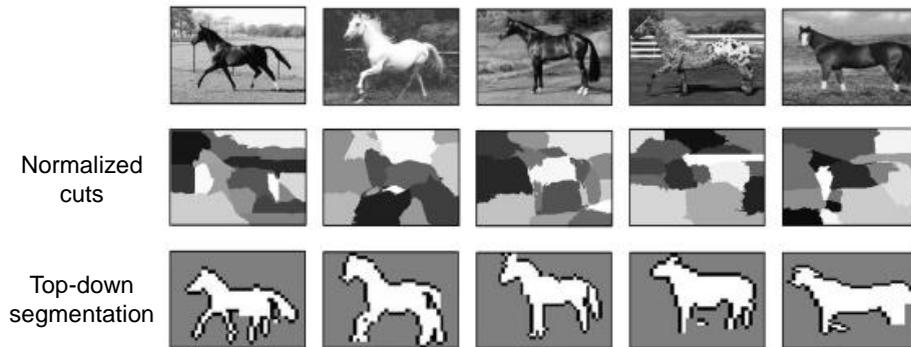
E. Borenstein and S. Ullman, [“Class-specific, top-down segmentation,”](#) ECCV 2002

A. Levin and Y. Weiss, [“Learning to Combine Bottom-Up and Top-Down Segmentation,”](#) ECCV 2006.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

Slide: Lazebnik 

Top-down segmentation



E. Borenstein and S. Ullman, "[Class-specific, top-down segmentation.](#)" ECCV 2002

A. Levin and Y. Weiss, "[Learning to Combine Bottom-Up and Top-Down Segmentation.](#)" ECCV 2006.

Computer Vision 2011-2012 - Prof. Sebastiano Battiato

