

IRS e DBMS multimediali

- ◆ Gestiscono diversi tipi di informazioni, potenzialmente rappresentate secondo diversi formati
 - immagini
 - audio
 - video
- ◆ Permettono di:
 - rappresentare dati corrispondenti a diverse tipologie di media
 - interrogare in modo uniforme dati con formati diversi
 - interrogare dati in formati diversi simultaneamente nel contesto della stessa query
 - recuperare gli oggetti dal supporto su cui risiedono, compatibilmente con il tipo di media che rappresentano

1

Problematiche

- ◆ La realizzazione di un sistema per il recupero di documenti multimediali deve affrontare le seguenti problematiche
 - rappresentazione dei documenti
 - misura della similarità
 - metodi di accesso e di ricerca

2

Caratteristiche generali

- ◆ Tipo di informazione
 - tipicamente non strutturata
 - un oggetto multimediale in genere è composto da diversi sotto-oggetti, ciascuno relativo ad un particolare media
- ◆ Interrogazione
 - spesso si deve analizzare il contenuto degli oggetti
 - come rappresentare il contenuto di un documento multimediale?
- ◆ Approccio tipico: uso di metadati associati ai documenti, che ne descrivono caratteristiche generali e contenuto
 - attributi descrittivi
 - features estratte dai documenti

3

Esempio



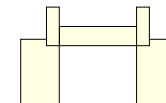
Attributi descrittivi

nome: Villa Medici

località: Roma

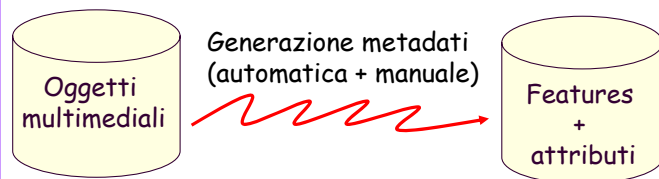
Features:

forma:



4

Rappresentazione con metadati



I sistemi mettono a disposizione funzionalità per estrarre feature e per utilizzare tali feature nel contesto delle interrogazioni

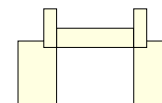
5

Rappresentazione con metadati

◆ Conseguenza

- i documenti multimediali potranno essere confrontati solo rispetto agli attributi e alle feature
- feature uguali non sempre si riferiscono ad oggetti uguali

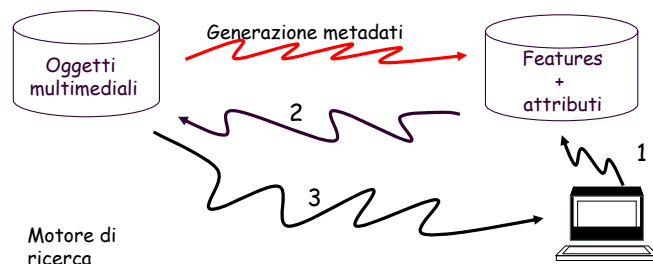
■ Esempio



si può riferire sia ad una villa che ad un'industria

6

Interrogazione con metadati



1. La query viene eseguita sui metadati (attributi + features)
2. dai metadati si risale ai documenti originali
3. i documenti originali vengono restituiti all'utente

7

Features per immagini: shape

◆ Forme (shape)

- l'idea è quella di estrarre delle forme dall'immagine
- possono rappresentare contorni o approssimazione di contorni



contorno

Approssimazione
contorno

8

Features semantiche

- ◆ un elemento significativo dal punto di vista dell'applicazione
 - Es: casa, viso, strada
 - se tali elementi sono chiaramente identificabili e possono essere facilmente riconosciuti, il ritrovamento può essere basato su di essi
- ◆ Problemi
 - L'estrazione di tali feature tipicamente richiede intervento umano
 - approccio limitato dalle tecniche esistenti di analisi delle immagini

9

Features: colore

- ◆ proprietà globale che non richiede conoscenza degli oggetti contenuti nelle immagini
- ◆ può essere determinato in modo automatico quindi è molto utilizzato
- ◆ spesso si usano istogrammi che rappresentano la composizione di colori in un'immagine
 - ogni componente dell'istogramma corrisponde ad un colore (256 o 64 componenti)
 - per rappresentare un'immagine, si associa ad ogni componente il numero di pixel dell'immagine più simili al colore considerato
 - il colore di ogni pixel in genere viene rappresentato come una tripla (R,G,B), dove R indica la percentuale di Rosso, G la percentuale di Giallo e B la percentuale di Blu

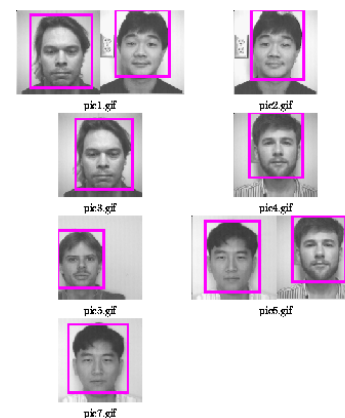
10

Una possibile vista logica per le immagini

- ◆ In generale, per rappresentare in modo sufficientemente dettagliato un'immagine è necessario utilizzare più features
- ◆ In generale, ogni immagine può essere descritta come:
 - proprietà globali
 - un insieme di elementi interessanti, ciascuno caratterizzato da un descrittore di forma, che rappresenta la shape dell'oggetto e/o la zona dell'immagine nella quale l'oggetto è collocato
 - proprietà locali
 - un descrittore di proprietà, che descrive le proprietà di un insieme di pixel nell'immagine (RGB, livelli di grigio per immagini in bianco e nero)

11

Esempio



12

Esempio

- ◆ Si consideri l'immagine pic1.gif
- ◆ l'immagine contiene due oggetti di interesse - o1 e o2 -
 - le shape di questi oggetti sono rappresentate dai rettangoli presenti nella figura
 - il descrittore di proprietà associato ad un insieme di pixel potrebbe avere la seguente forma
 - Red = 5
 - Green = 1
 - Blue = 3

13

Descrittori di proprietà

- ◆ Le proprietà vengono in genere associate ad insiemi di pixel
- ◆ In genere, ogni immagine viene associata ad una coppia di interi positivi (m,n) , chiamata la griglia di risoluzione dell'immagine
- ◆ una griglia (m,n) divide l'immagine in $(m \times n)$ celle di uguale dimensione, chiamate griglia dell'immagine
- ◆ ogni cella è costituita da un insieme di pixel
- ◆ le proprietà possono quindi essere associate a ciascuna cella
- ◆ le proprietà di ogni singola cella possono essere rappresentate come un vettore, in cui ciascuna componente corrisponde ad una determinata proprietà

14

Osservazione

- ◆ Le shape si possono interpretare come proprietà globali dell'immagine
- ◆ le proprietà sono invece locali a ciascuna singola cella
- ◆ dalle proprietà di ogni singola cella si possono ovviamente inferire le proprietà delle celle che rappresentano i vari oggetti identificati

15

Definizione di base di dati di immagini

- ◆ Una base di dati di immagini è una tripla $(GI, Prop, Rec)$ dove:
 - GI è un insieme di immagini a cui è stata associata una griglia
 - ogni immagine si può quindi vedere come una terna $(Immagine, m, n)$, dove (m, n) rappresenta la griglia di risoluzione
 - Prop è un insieme di proprietà di celle
 - Rec è una funzione che associa ad ogni immagine un insieme di shape, in base ad una qualche rappresentazione
- ◆ Prop rappresenta le proprietà locali delle immagini
- ◆ Rec rappresenta le proprietà globali delle immagini

16

Query per similitudine

- ◆ Le query possono coinvolgere sia proprietà locali che proprietà globali
- ◆ problema: come determino la similitudine tra due immagini?
- ◆ Approccio fondamentale: approccio metrico
 - si assume che esista una distanza metrica d con la quale confrontare ogni coppia di oggetti
 - più bassa è la distanza, più simili sono gli oggetti

17

Approccio metrico

- ◆ Una funzione da un insieme X a $[0,1]$ è una funzione distanza se soddisfa i seguenti assiomi per ogni x,y,z in X :
 - $d(x,y) = d(y,x)$
 - $d(x,y) \leq d(x,z) + d(z,y)$
 - $d(x,x) = 0$
- ◆ per ogni proprietà associata ad un'immagine può essere definita una funzione distanza
- ◆ la distanza tra due immagini si ottiene quindi combinando le distanze tra le varie proprietà (locali e/o globali), ad esempio applicando la distanza Euclidea

18

Esempio applicato alle proprietà globali

- ◆ Si consideri un insieme Obj di immagini (256×256). Si supponga che ogni cella sia associata a tre attributi (red, green, blue) che assumono un valore in $\{0, \dots, 8\}$
- ◆ un esempio di funzioni distanza tra due immagini $o1$ e $o2$ rispetto alle tre proprietà considerate è il seguente:
 - $diff_r[i,j] = (o1[i,j].red - o2[i,j].red)^2$
 - $diff_g[i,j] = (o1[i,j].green - o2[i,j].green)^2$
 - $diff_b[i,j] = (o1[i,j].blue - o2[i,j].blue)^2$
 - $d(o1,o2) = (\sum_{i=1,256} \sum_{j=1,256} (diff_r[i,j] + diff_g[i,j] + diff_b[i,j]))^{1/2}$

19

Approccio metrico per proprietà locali

- ◆ La distanza precedente rappresenta la distanza Euclidea tra i due oggetti
- ◆ se si considerano n proprietà locali e $t = h \times k$ celle, in base all'approccio metrico ogni immagine è vista come un insieme di m punti n -dimensionali
 - $f(l) = \{P_1, \dots, P_t\}$ $P_i = (x_{i1}, \dots, x_{in})$
- ◆ per determinare se le immagini l_1 e l_2 sono simili in base alla distanza Euclidea:
 - $f(l_1) = \{P_1, \dots, P_t\}$ $P_i = (x_{i1}, \dots, x_{in})$
 - $f(l_2) = \{Q_1, \dots, Q_t\}$ $Q_i = (y_{i1}, \dots, y_{in})$
 - $d(l_1, l_2) = (\sum_{i=1,t} d'(P_i, Q_i))^{1/2}$
 - $d'(P_i, Q_i) = (x_{i1} - y_{i1})^2 + \dots + (x_{in} - y_{in})^2$

20

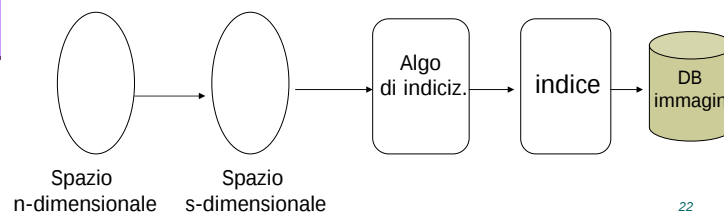
Approccio metrico per proprietà locali

- ◆ L'approccio metrico richiede che per ogni proprietà locale venga calcolata la distanza tra celle corrispondenti in immagini distinte
- ◆ ogni immagine è quindi interpretata come un insieme di punti
 - ogni punto è dato dai valori per le proprietà locali di una certa cella
 - se le proprietà sono n , i punti sono n -dimensionali
 - ogni immagine sarà caratterizzata da tanti punti quante sono le celle
- ◆ poiché il numero delle celle spesso è alto esistono approcci (che non vediamo) per mappare le immagini NON in un insieme di punti MA in un singolo punto in uno spazio s -dimensionale, con $s \leq n$
- ◆ Esempio: consideriamo la media degli n punti
 - in questo caso s coincide con n
- ◆ si noti che il punto ottenuto si può interpretare come una proprietà globale dell'immagine

21

Indicizzazione per proprietà locali

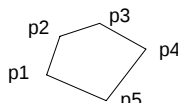
- ◆ I punti ottenuti possono essere utilizzati come base per definire le tecniche di indicizzazione
- ◆ nel caso limite in cui $s = 1$, si potrebbero usare I B-tree
- ◆ nel caso in cui $s > 1$, è necessario utilizzare specifici indici multidimensionali
- ◆ questi indici sono stati definiti nel contesto delle basi di dati spaziali



22

Approccio metrico per proprietà globali

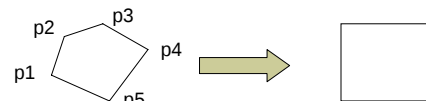
- ◆ Anche per gli oggetti (regioni) associati ad un'immagine è possibile applicare un approccio metrico
- ◆ in questo caso la distanza deve essere definita sulle shape
- ◆ anche le shape possono essere rappresentate come un insieme di punti
- ◆ Esempio
 - se la shape è rappresentata da una spezzata costituita da n segmenti, la spezzata può essere descritta da una sequenza di n punti bidimensionali



23

Approccio metrico per proprietà globali

- ◆ In questo caso è possibile passare da n punti a 2 punti considerando il minimo rettangolo che contiene la shape (minimum bounding box - MBB)
- ◆ 2 punti si possono vedere come un punto in uno spazio 4-dimensionale ($s = 4$)



24