

Sistemi Multimediali

Cenni di Retrieval di dati Multimediali

Ombretta Gaggi
Università di Padova

Problema

La ricerca su una base di dati testuale è un problema studiato a fondo al quale sono state trovate ottime soluzioni

- ricerca di parole chiavi, sinonimi, eliminazione delle stop words, ecc.
- i motori di ricerca sono un esempio di come questo problema possa essere risolto



I dati multimediali introducono molti problemi in più rispetto al testo, che rendono la ricerca più difficile fin dalla sua formulazione

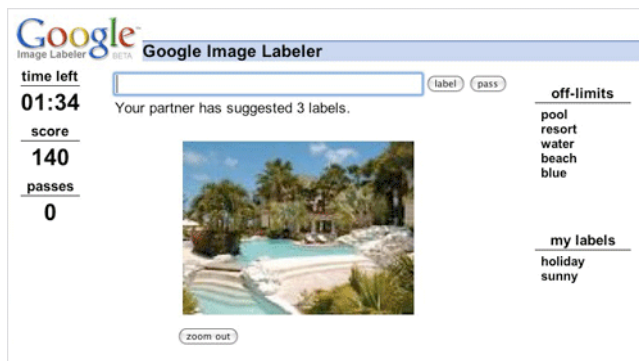
- i dati sono multidimensionali
- l'estrazione delle feature può essere computazionalmente costosa e difficile
- non esiste un modo predefinito di presentare l'interfaccia per la formulazione di una query, né per ritornare i risultati



Sistemi Ipermediali - 2

La soluzione semplice

La ricerca testuale può essere una soluzione se il database multimediale è completamente descritto da tag testuali



Sistemi Ipermediali - 3

Problemi

Ci sono immagini per cui non è semplice trovare dei tag utili per la ricerca



Sistemi Ipermediali - 4

Utilizzo di feature

Molti sistemi di retrieval di dati multimediali, in particolare immagini, si basano sull'estrazione di feature (ovvero caratteristiche basate su informazioni statistiche o spaziali dell'immagine)

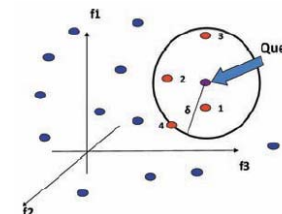
Le feature più comuni sono:

- **Istogramma dei colori**: un array tridimensionale (RGB) che conta i pixel di un particolare colore
- **Schizzo** (color layout)
- **Texture**: insieme di descrittori della texture, tipicamente basata sulle caratteristiche (*edge*) di un'immagine



Query

Ad ogni immagine viene associato un vettore $(f_1, f_2, f_3, \dots, f_n)$ di feature estratte. Ogni immagine diventa quindi un punto in uno spazio n -dimensionale.



Candan & Sapino, Data Management for Multimedia Retrieval 2010

Una query definisce un punto nello spazio e cerca gli elementi vicini (in un intorno). La vicinanza viene calcolata con una funzione di **similarità** che varia da sistema a sistema

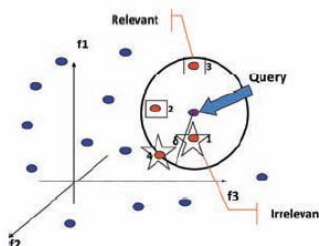


Relevance feedback

La definizione del punto da cercare non è un argomento banale nella creazione dell'interfaccia utente.

Le feature estratte potrebbero non rappresentare bene la caratteristica cercata

Il relevance feedback coinvolge l'utente per raffinare il risultato di una query che diventa quindi un processo iterativo



Candan & Sapino, Data Management for Multimedia Retrieval 2010



Problema del Dominio

Nel campo del retrieval di dati multimediali, il dominio ha un forte influenza

Si possono ottenere risultati completamente diversi nel caso in cui il dominio si

- **ampio** (es. un database di immagini per i siti che vendono fotografie e icone)
- **stretto** (es. un database di ecografie relative ad un certa ricerca sperimentale)



Problemi

Ogni sistema di multimedia retrieval soffre di due problemi:

1. Gap sensoriale: il divario presenta tra un oggetto nel mondo reale e la descrizione che se ne può trarre analizzando computazionalmente una sua registrazione
2. Gap semantico: il divario presente tra le informazioni che possono essere estratte da un dato visuale e l'interpretazione dello stesso dato da parte di un utente in una certa situazione



Istogramma di colori

È una feature di basso livello abbastanza semplice, non è influenzata dall'orientamento dell'immagine e dalla parziale occlusione degli oggetti

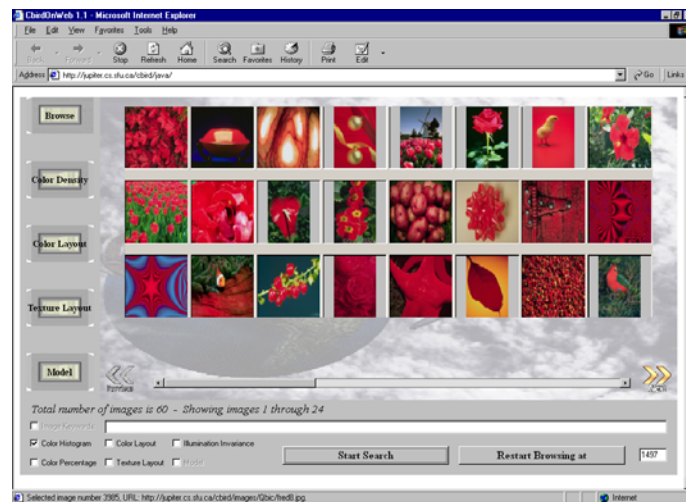
Questa feature viene calcolata quando l'immagine viene inserita nel database.

```
int hist[256][256][256];
for i=0..(N-1)
    for j=0..(M-1){
        R = image[i][j].red;
        G = image[i][j].green;
        B = image[i][j].blue;
        hist[R][G][B]++;
    }
```

Di solito si usano 3 bit per ogni le componenti rossa e verde e 2 bit per la componente blu nella memorizzazione dell'istogramma



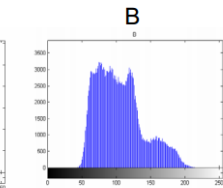
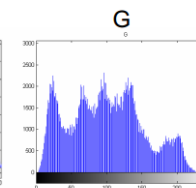
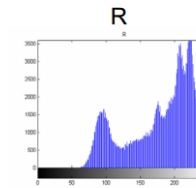
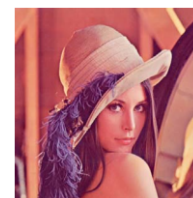
Risultati per una ricerca per istogramma di colori



Li & Drew, Fundamentals of Multimedia, 2003



Istogramma di colori



Histogram Intersection

È una misura di similarità tra gli istogrammi, veloce da calcolare ma sensibile alla quantizzazione dei colori

L'istogramma viene normalizzato per essere indipendente dalla dimensione dell'immagine

L'intersezione è definita come

$$intersection = \sum_{j=1}^n \min(H_i^j, H_m^j)$$

dove j è un colore

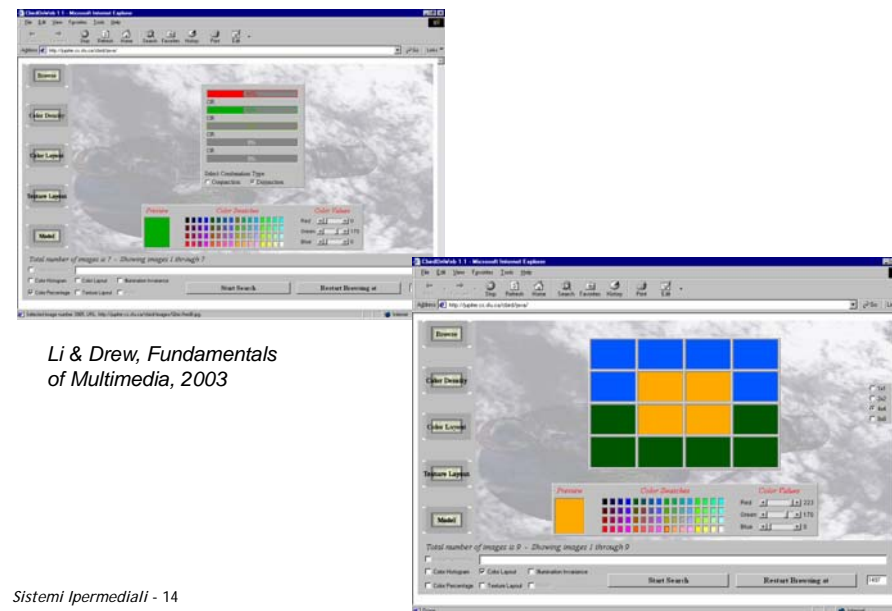
H_i è l'istogramma di un'immagine nel db

H_m è l'istogramma scelto per la ricerca

Più il valore di intersezione è vicino a 1, più le immagini sono simili



Color density and color layout



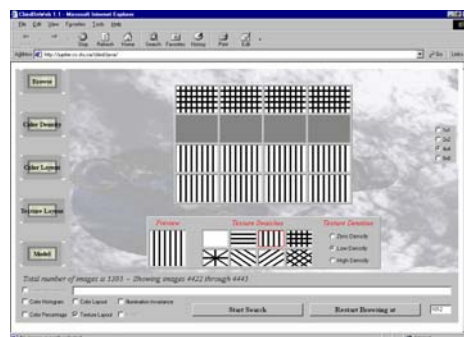
Li & Drew, Fundamentals of Multimedia, 2003

Texture

Il termine *texture*, in italiano tessitura, rappresenta il dettaglio di una superficie

La feature *texture layout* è simile alla feature *color layout* e permette di descrivere la distribuzione di texture desiderata

Per una texture possono definire la densità, e la direzione



Esempi di texture

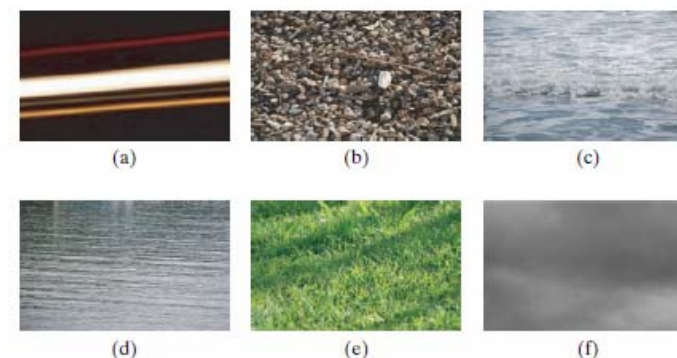


Figure 2.11. (a) A relatively smooth and directional texture; (b) a coarse and granular texture; (c) an irregular but fractal-like (with elements self-repeating at different scales) texture; (d) a regular, nonsmooth, periodic texture; (e) a regular, repeating texture with directional elements; and (f) a relatively smooth and uniform texture.

Candan & Sapino, Data Management for Multimedia Retrieval 2010



Confronto tra texture

Per confrontare due texture tra loro si usa il *texture histogram* ovvero una matrice bidimensionale che calcola la *direzionalità* e la *edge separation* per ogni immagine

La direzionalità misura l'orientamento di un edge e la edge separation determina la distanza tra due edge paralleli

Per estrarre la *edge map*, il calcolo viene limitato alla sola componente Y



Figure 2.13. (a) Mountain ridges commonly have self-repeating triangular shapes. (b) This is a fragment of the texture in Figure 2.11(c).

Candan & Sapino, *Data Management for Multimedia Retrieval* 2010



Invarianza per illuminazione

Una diversa illuminazione di uno stesso oggetto può introdurre molte differenze nelle immagini catturate

Per evitare questo problema, tutte le componenti colori vengono normalizzate, poi viene calcolato un istogramma dei colori per la *cromaticità*

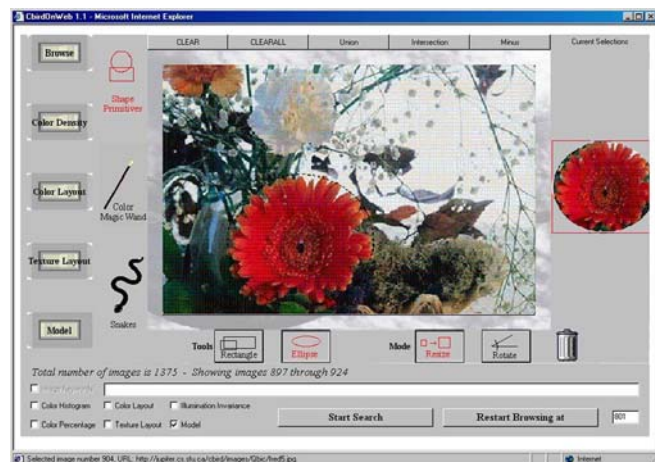
$$\{R, G\} / (R + G + B)$$

L'istogramma bidimensionale viene poi trattato come un'immagine, compresso usando DCT e presi i primi 36 valori per ridurre lo spazio di memorizzazione

La misura di similarità viene fatta nel dominio compresso, utilizzando la distanza euclidea tra due vettori di 36 componenti



Model-based object search



Li & Drew, *Fundamentals of Multimedia*, 2003



Oggetti con forme evidenti

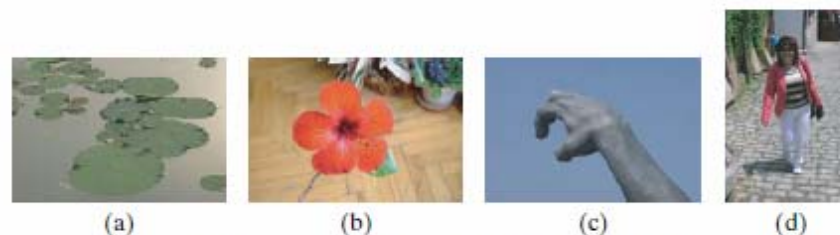


Figure 2.16. Sample images with dominant shapes.

Candan & Sapino, *Data Management for Multimedia Retrieval* 2010



Model-based object search

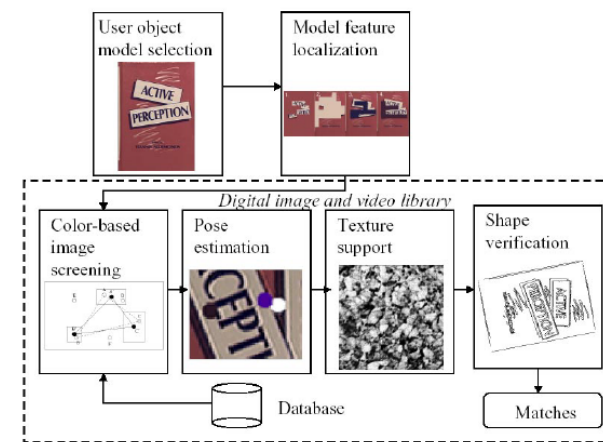
È un caso di query-by-example

L'utente può selezionare una porzione di immagine utilizzando forme semplici: cerchi, quadrati, ellissi, ecc.

Una volta selezionata la parte desiderata, vengono calcolate le feature di quella porzione di immagine (es. istogramma dei colori, cromaticità, ecc.)

Un passo successivo considera scala, rotazione e traslazione dell'oggetto all'interno di altre immagini

Model-based object search



Li & Drew, Fundamentals of Multimedia, 2003



Ricerca di video



(a)



(b)

Fig. 18.19: Digital video and associated keyframes, *beach* video.
(a): Frames from a digital video. (b): Keyframes selected.
Li & Drew, Fundamentals of Multimedia, 2003



That's all folk!

...Grazie per l'attenzione!

