

Sistemi per il recupero delle informazioni

Gabriele Pozzani

A.A. 2013/2014

**Corso di Laurea Magistrale in
Editoria e Giornalismo**

Sistemi per il recupero delle informazioni multimediali

Multimedia

- Un'informazione multimediale è una qualunque informazione digitale composta di diverse parti
 - Testo
 - Audio
 - Video
 - ...
- Quindi documenti non solo testuali
- Solitamente poco o non strutturata

3

Recupero delle informazioni multimediali

- Consiste nel recupero di testo, audio, video in risposta ad un'interrogazione di un utente e il loro ordinamento in base ad un qualche grado di rilevanza
- Come può avvenire una ricerca multimediale?
 - Descrivendo il suo contenuto
 - tecnicamente, simbolicamente, semanticamente (?)
 - La semantica è esterna al dato e la sua definizione dipende dal media
 - tutta l'immagine, l'audio o alcuni componenti
 - Fornendo un esempio dell'informazione desiderata
 - Immagine/audio simile
 - proprietà simili (sketch)
 - Attraverso i metadati

4

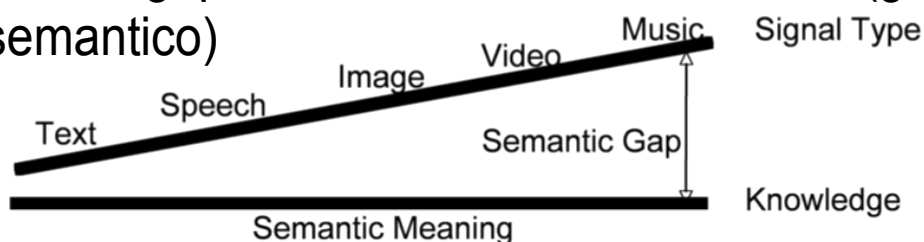
Caratteristiche dei SRI multimediali

- Informazioni
 - di natura semplice ma di significato complesso
 - il concetto di struttura è variabile e dipende dal media (immagine, filmato, audio)
 - elevate dimensioni per ogni esemplare di dato
- Interrogazioni
 - scarsa corrispondenza tra forma e significato
 - specifiche "sintattiche", ricerca per somiglianza
 - iterazioni successive, analisi di rilevanza
- Aggiornamenti
 - non frequenti, generalmente fuori linea
- Problemi specifici
 - la rappresentazione codificata non contiene il significato del dato
 - la codifica e la rappresentazione influiscono sui sistemi di gestione

5

Semantica

- Il principale problema con i dati multimediali è l'associazione ad essi di una semantica
 - i dati multimediali hanno un contenuto informativo molto ricco ma parte della informazione non è esplicita e può essere estratta solo analizzando un contesto più ampio
 - Vi è un gap tra il dato e la sua semantica (gap semantico)



6

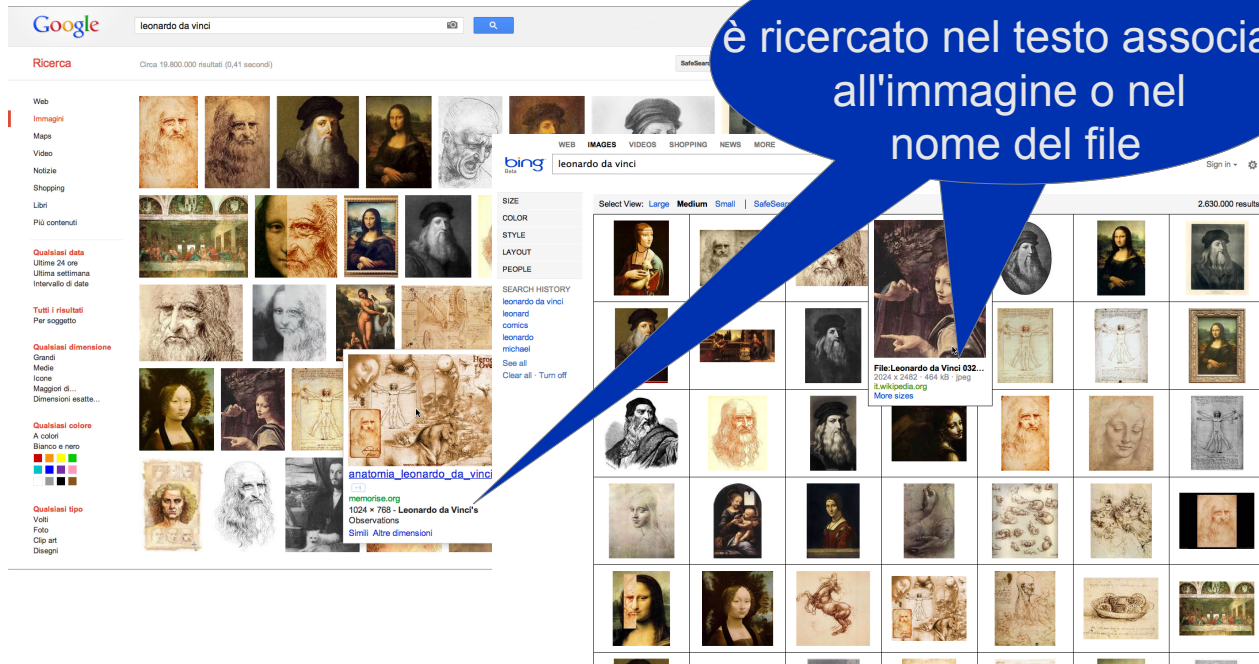
Gap semantico

- Un problema fondamentale è il riconoscimento degli oggetti in immagini, video, e audio
 - Non solo trovarli, ma anche interpretarli!!!
 - Una persona può guardare un'immagine e identificare facilmente visi ed altri oggetti qualunque sia la loro forma, colore, ecc...
 - Alcuni contenuti sono soggettivi ed “emozionali”
 - La velocità e il tono di un parlato
 - L'espressione di un viso
 - È un problema in generale ancora irrisolto

7

Come la ricerca di immagini funziona (I)

- Quando, ad esempio, Google trova una nuova immagine cerca di classificarla (e di comprenderne il contenuto) analizzando:
 - Il nome del file
 - Metadati inclusi nel file (e.g., metadati Exif)
 - La sua didascalia/titolo
 - Il testo della pagina in cui l'immagine si trova
- Utilizza queste informazioni per taggare l'immagine con termini di ricerca
- L'uso del testo della pagina porta anche ad associare un'immagine con termini che in realtà riguardano altre immagini nella stessa pagina
 - Provate ad esempio a cercare "Massimo Bellotto" su Google Image

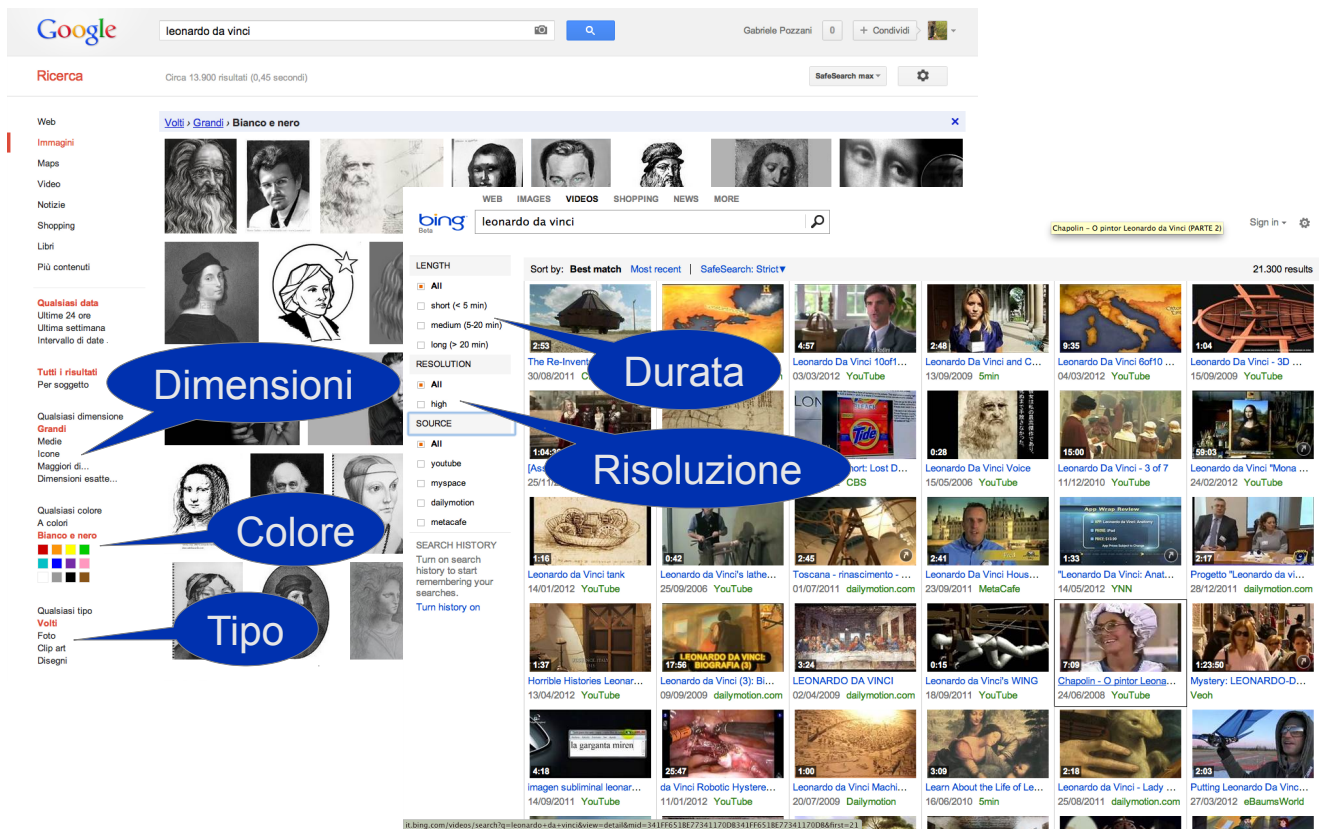


9

Come la ricerca di immagini funziona (II)

- In secondo luogo le immagini vengono analizzate in alcune loro caratteristiche “tecniche”
 - Dimensioni
 - Colore/BN
 - Colore principale (attraverso l'analisi dell'istogramma dell'immagine e altre tecniche)

10



11

Come la ricerca di immagini funziona (III)

- Infine si cerca una classificazione automatica delle immagini e del loro contenuto (senza l'uso di testi “di contorno”)
 - Si basa su tecniche quali il “machine learning” e le “reti neurali”
 - Ad esempio Google ha acquisito a maggio 2013 una start-up dell'Università di Toronto che lavora proprio in questo ambito
- L'analisi e classificazione del contenuto delle immagini si va ad “integrare” con la “ricerca semantica”

12

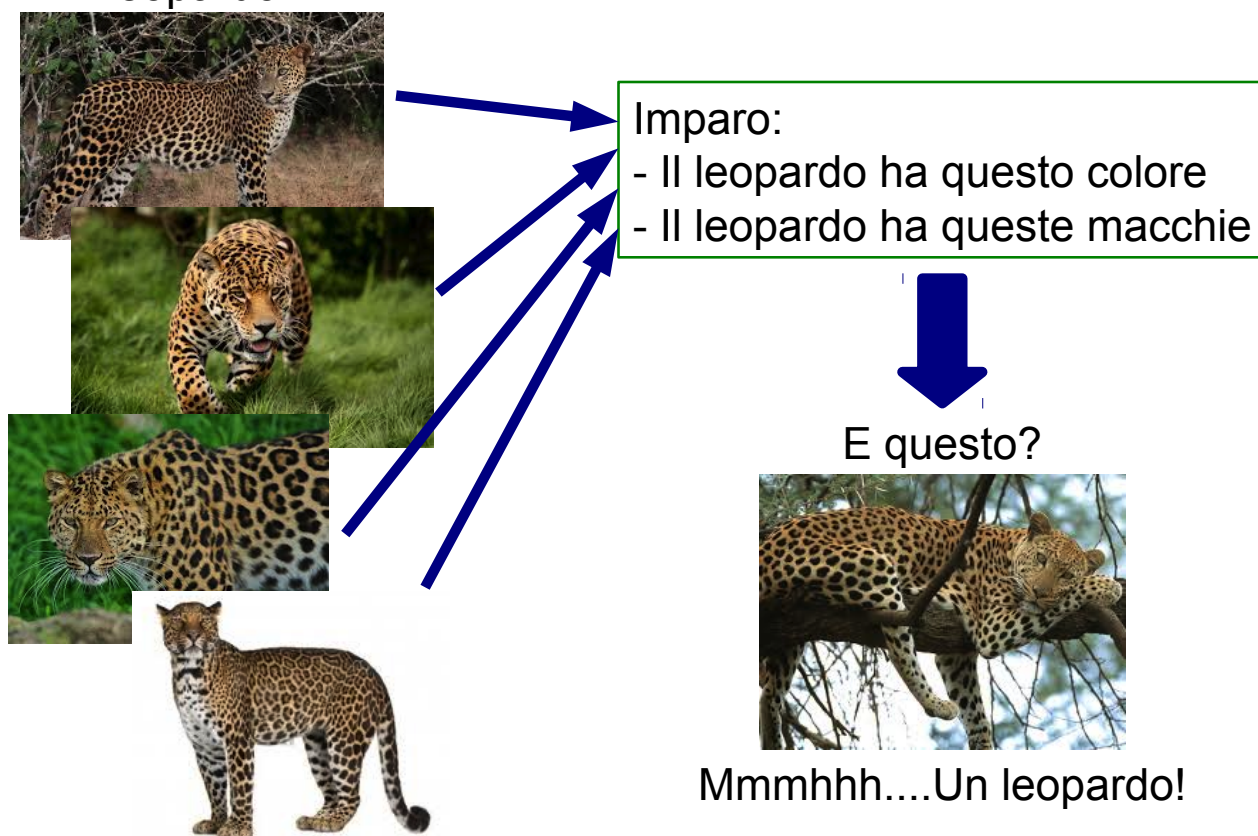
Come la ricerca di immagini funziona (IV)

- L'idea è di creare un sw che
 - Parta da delle immagini “di riferimento” che rappresentano determinati soggetti
 - Impari quali sono le caratteristiche fondamentali delle immagini per ogni soggetto
 - In modo che nuove immagini possano eventualmente essere riconosciute come rappresentanti uno dei soggetti

13

Come la ricerca di immagini funziona (V)

Leopardo



14

Come la ricerca di immagini funziona (VI)

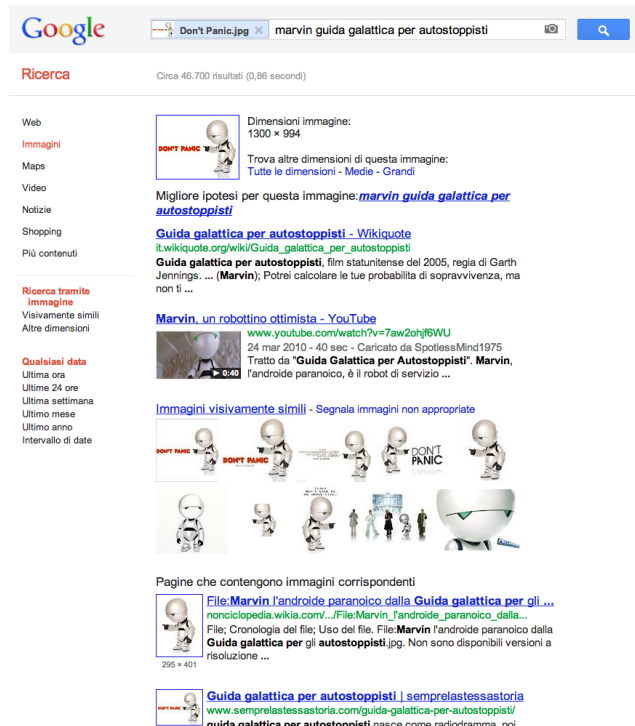
- Ovviamente la varietà con cui uno stesso soggetto può presentarsi rende MOLTO difficile il riconoscerlo
 - Facile sbagliare
- Ma la ricerca avanza e le tecniche si affinano
- Queste ultime tecnologie sono attualmente disponibili nella ricerca tra le foto caricate sul proprio profilo Google+ (ma non è detto funzioni bene ;))

<http://googleresearch.blogspot.it/2013/06/improving-photo-search-step-across.html>

Content-based image retrieval (I)

- Ricerca di immagini basandosi su alcune caratteristiche “tecniche” del contenuto
 - Non il contenuto semantico
 - Ma le caratteristiche del contenuto
 - Colore e sua distribuzione
 - Forma
 - Luminosità
 - Contorni
 - L'analisi di queste caratteristiche permette anche di
 - Catalogare le parti di un'immagine (sfondo, primo-piano, ...)
 - Identificare e taggare singoli oggetti perché se ne conoscono le proprietà fondamentali
- Query-by-example
 - Viene fornita un'immagine di riferimento e si cercano immagini “simili”

Content-based image retrieval (II)

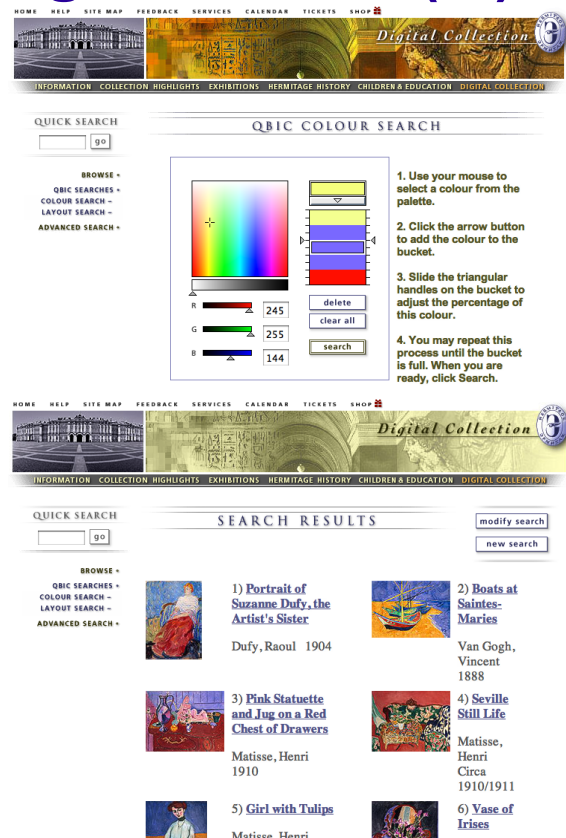


- Google permette il query-by-example
- Analizza e ricerca
 - Colore
 - Linee
 - Punti di riferimento
 - Texture
- Ritorna immagini con caratteristiche simili
- Analizza le pagine in cui tali immagini si trovano per
 - associare una descrizione testuale usata per cercare documenti “interessanti”

17

Content-based image retrieval (III)

- QBIC
 - Sviluppato da IBM
 - Abbandonato
 - Usato dall'Hermitage di San Pietroburgo per la ricerca di opere d'arte sulla base di
 - Colore
 - Layout



18

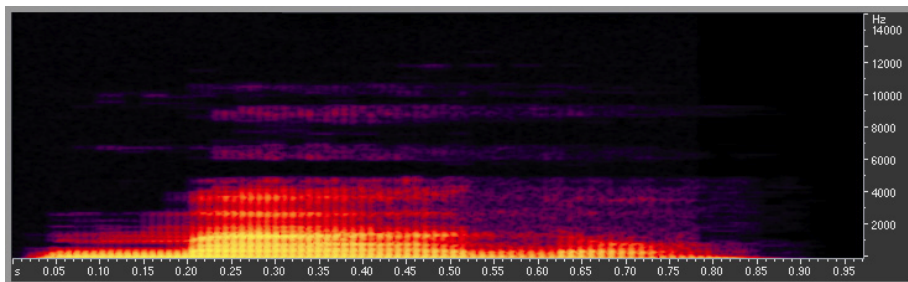
Recupero di audio e musica

- Recupero di tracce audio che risultino rilevanti rispetto ad una richiesta
 - Fingerprinting: ricerca di tracce audio a partire da un breve frammento audio (query-by-example, query-by-humming)
 - Riconoscimento del parlato: riconoscere il testo di una traccia audio contenente del parlato
 - Riconoscimento dello speaker: riconoscere le persone che parlano in una traccia audio
 - Recupero di tracce parlate: ricerca di tracce parlate rispetto ad una query fornita testualmente

19

Fingerprinting

- Applicazione di successo
- Basata sull'analisi dello spettrogramma delle tracce audio



- Complicata da eventuali rumori e distorsioni nelle tracce
- Il software più famoso è probabilmente Shazam
 - Applicazione per dispositivi mobili
 - A partire da un frammento audio di 10s ne costruisce il fingerprint e lo ricerca in DB centralizzato contenente i fingerprint di milioni di canzoni

20

Riconoscimento del parlato

- Il riconoscere le parole in una traccia audio
- Applicazioni tipiche e funzionanti sono
 - Apple Siri Personal Assistant
 - Google Voice/Now
 - Bing voice search
 - Microsoft Tellme for Windows Phone 7
- Una volta riconosciute le parole, queste possono essere usate per una classica ricerca o per eseguire azioni

21

Riconoscimento dello speaker

- Riconoscere chi sta parlando in una traccia senza riguardo a cosa sta dicendo
- Basato su diversi approcci matematici, tra cui
 - Speaker-dependent speech recognition: modellare e riconoscere le peculiarità nella pronuncia di ogni persona

22

Recupero di tracce parlate

- Recuperare le tracce audio che rispettano una query testuale
- Due diversi approcci basati su
 - Keyword
 - Riconoscere alcune parole chiave nell'audio
 - Limita l'utente ad usare un insieme limitato di termini nella query
 - Riconoscimento fonetico
 - Riconoscere il parlato a livello di fonemi
 - Complicato dalla similarità di alcuni fonemi