

Corso di Laurea Specialistica in
Informatica dei Sistemi Avanzati e dei Servizi di Rete

Studio introduttivo alle tecniche di Video Retrieval

Guido Vicino

Content Based Video Retrieval

- Che cos'è un sistema CBVR
- Possibili applicazioni di tali sistemi
- Problemi che insorgono nell'analisi di un documento video
- Interesse commerciale verso questi strumenti
- Com'è stata condotta l'analisi su questi oggetti

Come si analizza e recupera informazione da un video (1)

- Analisi delle similarità
 - Similarità per colore
 - Similarità per trama
 - Similarità per forma
- Segmentazione temporale
 - Fotogramma, shot, scena e transizioni
 - Flussi non compressi
 - Flussi compressi

Come si analizza e recupera informazione da un video (2)

- Analisi ed annotazione di un filmato:
 - Analisi del contenuto di scena
 - Analisi del movimento della telecamera
 - Analisi del movimento degli oggetti
 - Analisi della colonna sonora
- Accesso ai contenuti estratti:
 - Visualizzazione dei fotogrammi chiave
 - Formulazione delle interrogazioni

Analisi di similarità per colore (1)

- Studio del colore tramite aspetti:
 - fisici
 - neuro-fisiologici
 - psicologici
- Energia luminosa come lunghezza d'onda $E(\lambda)$
- Lunghezze d'onda percepite dall'uomo (RGB)

Analisi di similarità per colore (2)

- Analisi dell'energia spettrale $C(\lambda)$:

$$\alpha_i(C) = \int_{\lambda_{min}}^{\lambda_{max}} S_i(\lambda) C(\lambda) d\lambda$$

- Caratteristiche del colore percepito:
 - Brillanza
 - Chiarezza
 - Tinta
 - Saturazione

Analisi di similarità per colore (3)

- Studio dei colori tramite vettori n-dimensionali
 - RGB(Red, Green, Blue)
 - CMY(Cyan, Magenta, Yellow)
 - LUV(Luminanza e cromaticità)
- Studio tramite analisi sulla distribuzione del colore:

$$C_d = \sqrt{(L_q - L_t)^2 + 4(U_q - U_t)^2 + 4(V_q - V_t)^2}$$

Analisi di similarità per trama

- Utile per discriminare tra oggetti di colore simile ma dal disegno diverso
- Utilizzo di modelli statistici indipendenti dal colore, ad esempio utilizzando:
 - Ruvidezza
 - Contrasto
 - Orientamento

$$T_d = \sqrt{\frac{(\alpha_q - \alpha_t)^2}{\sigma_\alpha^2} + \frac{(\beta_q - \beta_t)^2}{\sigma_\beta^2} + \frac{(\theta_q - \theta_t)^2}{\sigma_\theta^2}}$$

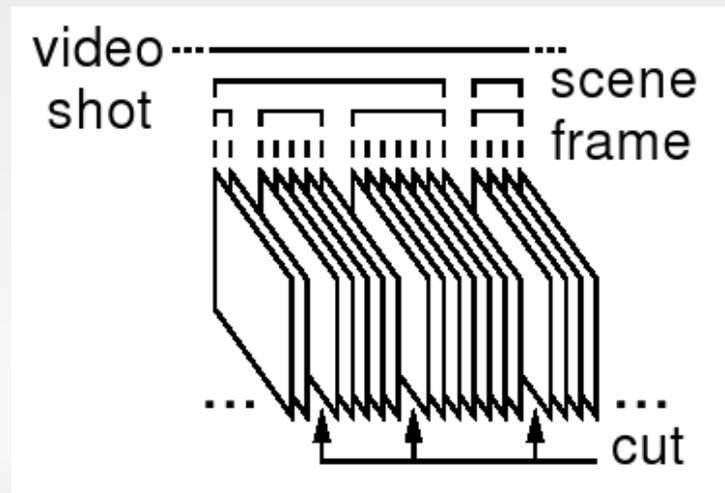
Analisi di similarità per forma

- Caratteristiche globali utilizzate:
 - area, allungamento, compattezza, direzione degli assi
- Caratteristiche locali utilizzate:
 - angoli, punti e segmenti caratteristici
- Studio delle trasformazioni geometriche tra un oggetto ed un altro:

$$T: A \rightarrow B$$

Segmentazione temporale (1)

- Si vuole suddividere un documento video in unità definite:
 - Frame
 - Shot
 - Scena



Segmentazione temporale (2)

- Fattori d'analisi relativi alle transizioni delle immagini:
 - Movimento degli oggetti e della telecamera
 - Luminosità e cambiamenti di luce
 - Cambiamento di shot
 - Bruschi
 - Progressivi
- Studio dei flussi non compressi
- Studio dei flussi compressi (MPEG)

Flussi non compressi (1)

- Rilevazione dei tagli (cut) tramite:
 - Istogrammi del colore tra immagini

$$\sum_j \frac{(H_{t+1}(j) - H_t(j))^2}{H_{t+1}(j) + H_t(j)}$$

- Confronto a blocchi

$$L_i = \frac{\left[\frac{\sigma_t^2 + \sigma_{t+1}^2}{2} + \left(\frac{\mu_t - \mu_{t+1}}{2} \right)^2 \right]}{\mu_t^2 \mu_{t+1}^2}$$

Flussi non compressi (2)

- Rilevazione dei tagli (cut) tramite:
 - Confronti basati su proprietà locali, quali analisi del cambiamento dei bordi degli oggetti e confronti sui pixel entranti/uscenti:

$$D(i, i+1) = \frac{\sum_{x=1}^X \sum_{y=1}^Y \sum_c |P_i(x, y, c) - P_{i+1}(x, y, c)|}{X * Y}$$

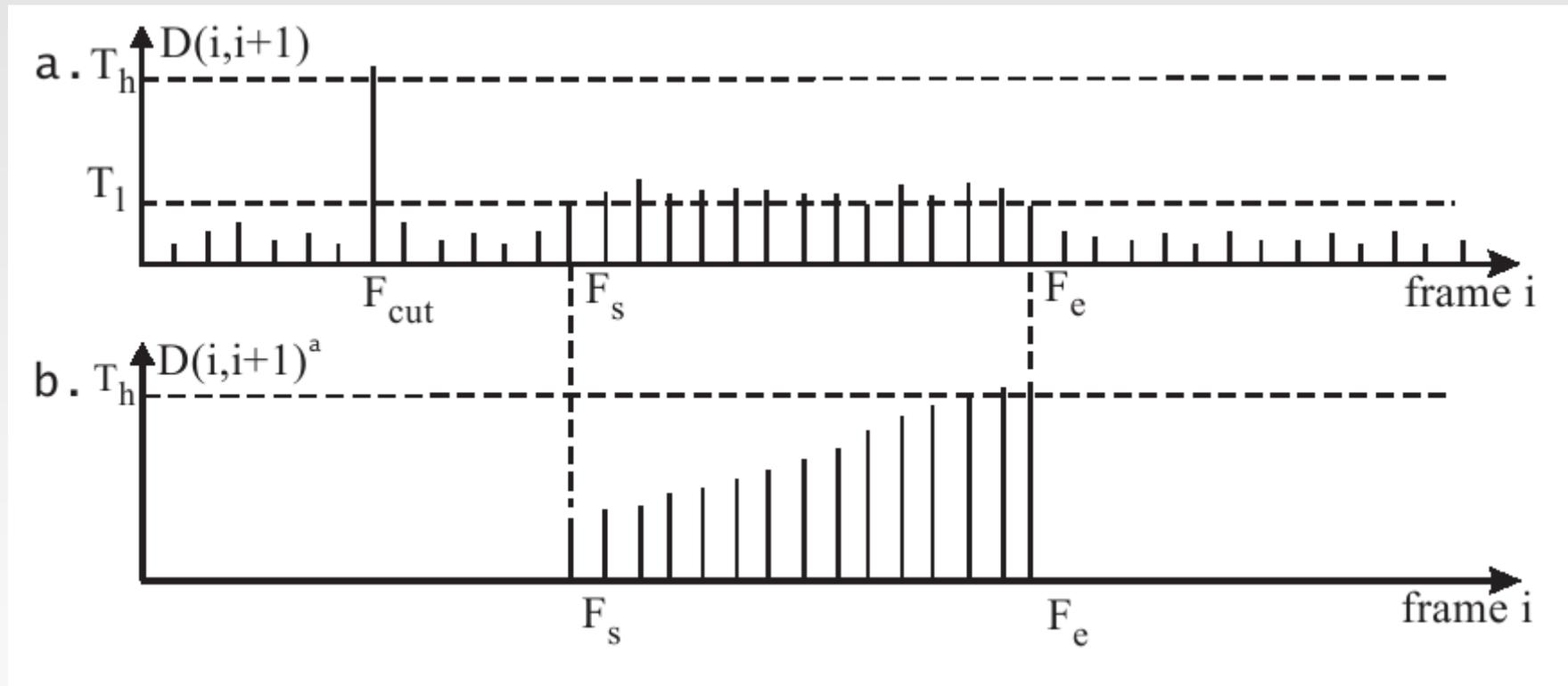
Flussi non compressi (3)

- Transazioni graduali:
 - Fade
 - Dissolve
 - Wipe



Flussi non compressi (4)

- Twin tresholding mechanism



Flussi compressi

- Sfruttano le proprietà interne del formato MPEG:
 - Utilizzo della compensazione di moto tramite vettori di movimento (*motion vector*)
 - Analisi dei coefficienti relativi alla *trasformata discreta del coseno (DCT)*

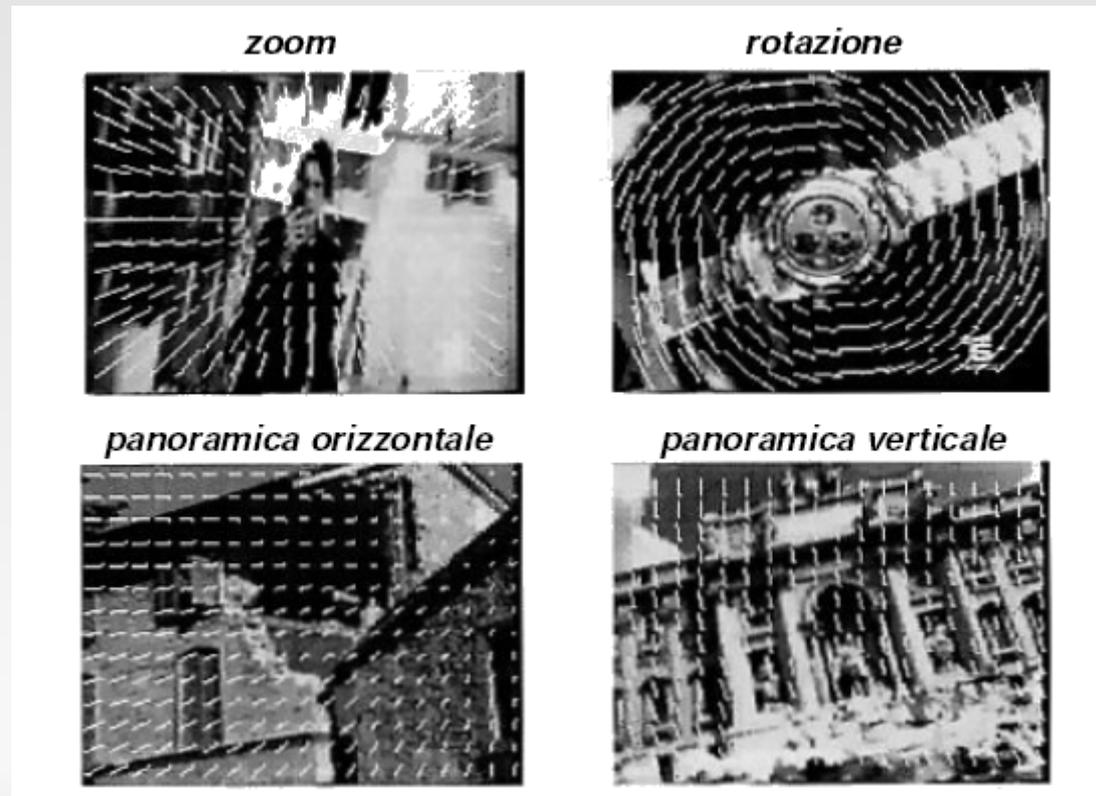
$$D_{cut} = 1 - \frac{\vec{C}_f \cdot \vec{C}_{f+k}}{|\vec{C}_f| |\vec{C}_{f+k}|}$$

Analisi del contenuto della scena

- Si limita il campo di studio a determinate tipologie di filmati:
 - Notiziari e telegiornali
 - Video sportivi
 - Video commerciali
 - Film
- L'annotazione ed indicizzazione viene fatta in maniera automatica ma assistita

Analisi del movimento della telecamera (1)

- Si vuole distinguere i vari movimenti della telecamera ai fini dell'indicizzazione:



Analisi del movimento della telecamera (2)

- Si rilevano i movimenti tramite lo studio del flusso compresso
- Si rilevano le operazioni di camera andando a creare soglie sulla differenza tra i vettori movimento θ_k e il vettore modale θ

$$\sum_N^k |\theta_k - \theta| \leq \Theta_p$$

Analisi della colonna sonora

- Fasi dell'analisi audio:
 - Rivelazione, segmentazione di voce, musica e suoni
 - Identificazione dello speaker nei sonori parlati
 - Trascrizione ed etichettatura del linguaggio naturale
- Come viene fatto?
 - Parametrizzazione dell'audio
 - Quantizzazione
 - Calcolo delle distanze

Annotazione tramite icone

- Si utilizzano icone per
 - Rappresentazione semantica
 - Rappresentazione temporale
- Evidenziare categorie
 - clima, condizioni di luce, locazioni, caratteri ed oggetti
- Evidenziare situazioni
 - Movimento del protagonista, di un oggetto

Fotogrammi chiave

- Si vuole fornire una descrizione globale del filmato:
 - Estrazione dei fotogrammi chiave
 - Come rappresentare?
 - Come ridurre opportunamente?
 - Costruzione di mosaici



Tipi di interrogazioni

- Esistono diversi tipi di interrogazioni:
 - Tramite Sketch
 - Query by Example (QBE)
 - Motion based
 - Testuali
- Sistemi reali provati:
 - VideoQ
 - YouTube, GoogleVideo

Conclusioni

- Diversa letteratura ma poche implementazioni aperte
- Diverse applicazioni nel campo dell'editing video:
 - Creazione automatica di trailer
 - Aumento dell'interesse dell'utente medio verso la ricerca di documenti video
- Interesse personale alle applicazioni legate alla video sorveglianza