

InformaticaUmanistica

Lezione 4

Gestione del video

Pasquale Savino

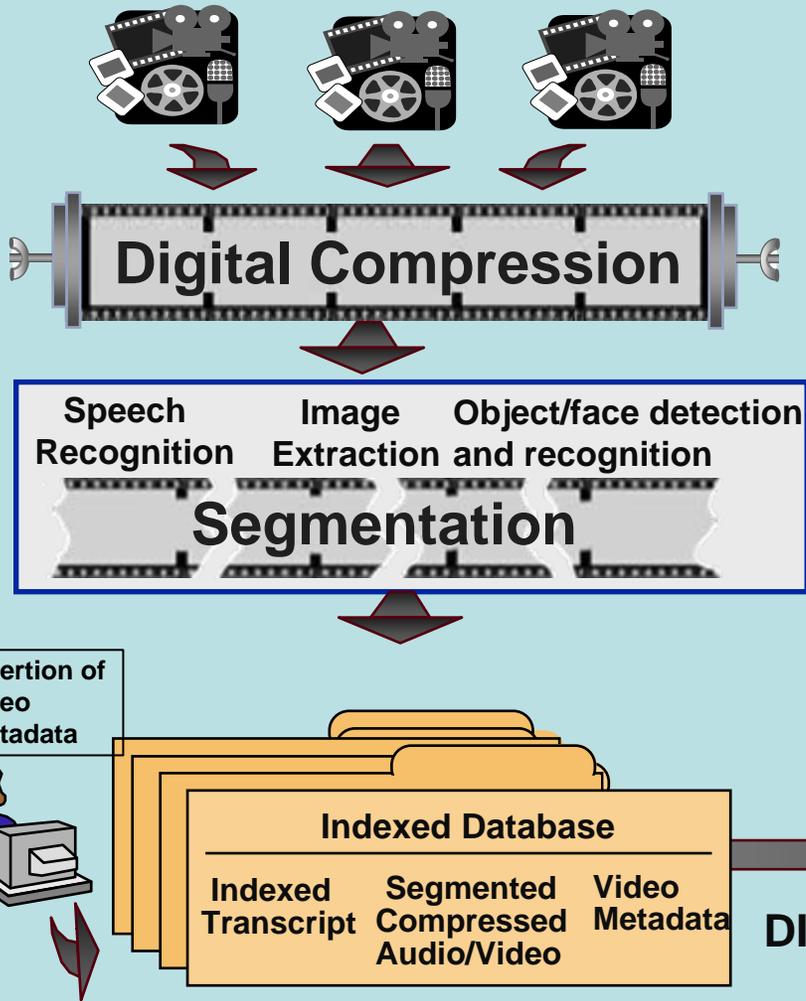
ISTI - CNR



UNIVERSITÀ DI PISA

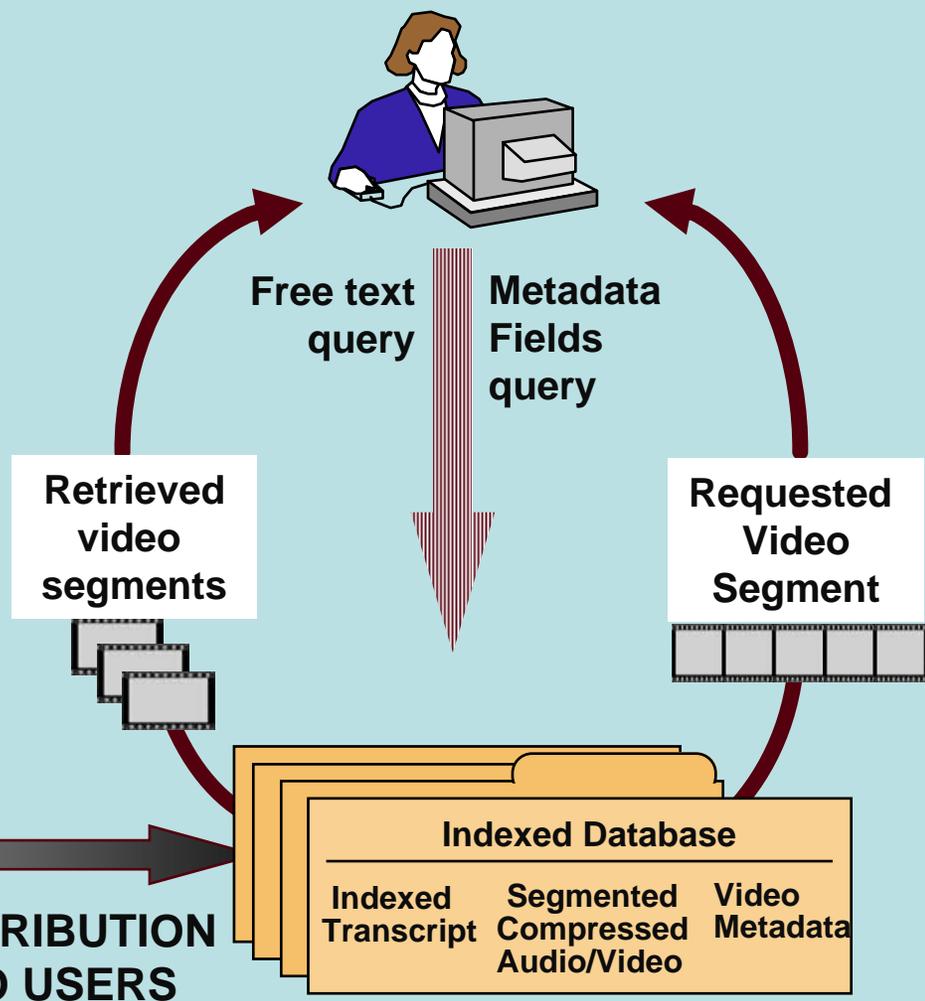
Library Creation

Offline



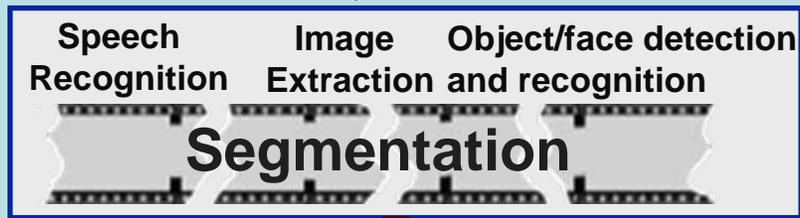
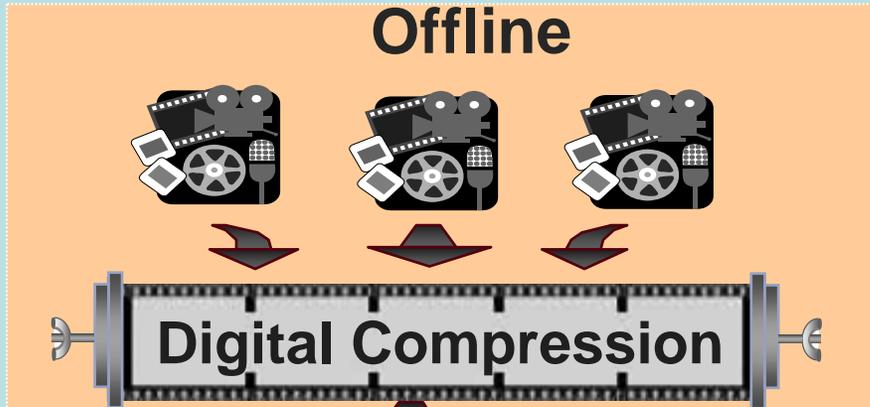
Library Exploration

Online

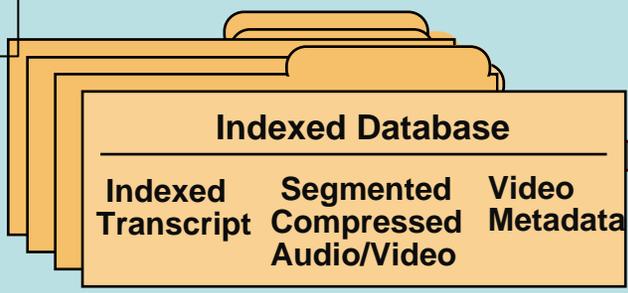


Library Creation

Offline



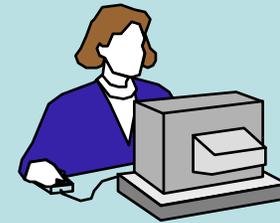
Insertion of video metadata



DISTRIBUTION TO USERS

Library Exploration

Online

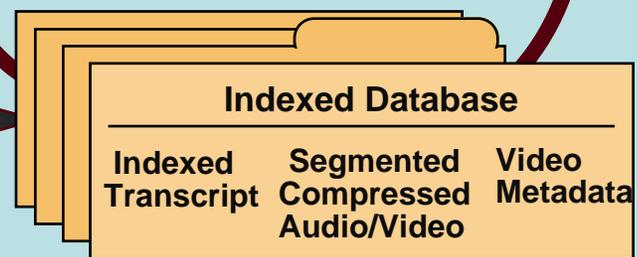


Free text query

Metadata Fields query

Retrieved video segments

Requested Video Segment



Tipi di dati presenti nel video

- ◆ **Un video digitale è composto di una sequenza di immagini (frame) che vengono visualizzate con un determinata frequenza, più (opzionalmente) una traccia audio.**
- ◆ **Il video può essere analizzato da diversi punti di vista, che possono essere anche combinati tra loro**
 - La parte audio può essere analizzata individuando le parti che contengono
 - **Parlato**
 - **Suono/musica**
 - Si può analizzare la sequenza dei frame, quindi l'evoluzione temporale del video e la sua organizzazione in shot e sequenze.
 - Si possono considerare i singoli frame come immagini
- ◆ **Da tutte queste componenti è possibile estrarre informazione da utilizzare per l'indicizzazione e per la ricerca del video.**

Caratteristiche del video digitale

- ◆ **Il video può essere descritto come una sequenza di frame visualizzati con una determinata frequenza, che varia a seconda della qualità del video e del formato di rappresentazione utilizzato.**
 - NTSC 30 frames/sec, PAL 25 f/s, HDTV 60 f/s
 - Nella maggior parte dei casi vi è una piccola variazione tra un frame ed il successivo.
- ◆ **Ogni frame è un'immagine con una risoluzione che dipende dalla qualità del video e dal formato**
 - 768 x 576 PAL, 720 x 480 NTSC
- ◆ **Da questi dati si deduce che un video non compresso occupa molto spazio su disco e richiede una elevata banda di comunicazione durante la trasmissione**
 - Per esempio, un secondo di video PAL non compresso richiede
 $768 \times 576 \times 16 \times 25 \sim 176 \text{ Mbit/s}$

Memorizzazione e trasmissione di video digitale [1/3]

- ◆ **La dimensione dei file video impone, in molti casi, di comprimere il video sia per memorizzarlo che per trasmetterlo.**
- ◆ **Nel caso del video si possono ottenere livelli di compressione molto alti, a causa delle caratteristiche del video stesso:**
 - **Correlazione spaziale:** in un singolo frame è possibile individuare una correlazione tra pixel adiacenti (come avviene nelle immagini), per cui è possibile utilizzare le tecniche di compressione delle immagini
 - **Correlazione temporale:** normalmente ogni frame differisce poco dai frame che lo seguono e lo precedono. Molto spesso il valore di un pixel rimane invariato per diversi frame successivi.
 - Una parte significativa del video non viene percepita da chi guarda il video.

Memorizzazione e trasmissione di video digitale [2/3]

- ◆ **Le tecniche di compressione si possono dividere in due grandi categorie**
 - **Lossless compression** (compressione senza perdita) che permette di comprimere e decomprimere il video senza alcun degrado della qualità
 - **La *lossless compression* permette di ottenere livelli di compressione bassi**
 - **Si ottiene ad esempio se ogni frame viene compresso in JPEG senza perdita. Questa tecnica di compressione si chiama MJPEG (Motion JPEG). In questo caso si sfrutta solo la correlazione spaziale.**
 - **Altri esempi di metodi di compressione senza perdita sono run-length coding e Huffman coding**

Memorizzazione e trasmissione di video digitale [3/3]

- **Lossy compression**, (compressione con perdita). In questo caso l'intero ciclo di compressione/decompressione introduce produce un video diverso dall'originale. L'obiettivo consiste nell'individuare delle tecniche di compressione che non introducono un degrado percepibile
 - **La compressione con perdita permette di ottenere livelli di compressione molto alti**
 - **Esempi di standard di compressione con perdita sono quelli della famiglia MPEG (MPEG1, MPEG2)**

MPEG

◆ MPEG (Moving Pictures Experts Groups)

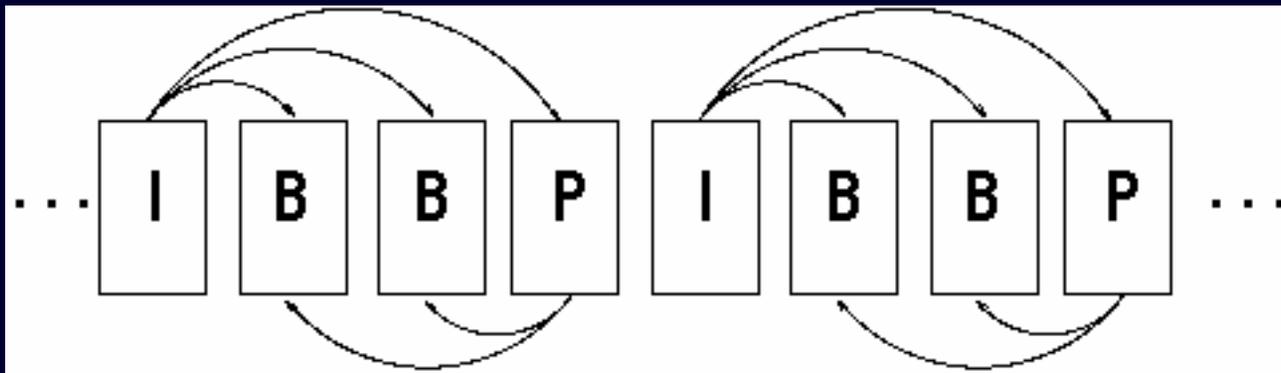
- MPEG1 ha un bit-rate di 1.5Mb/sec
 - Viene utilizzato per la compressione di video di qualità VHS. È stato sviluppato per la memorizzazione del video su CD-ROM.
 - Risoluzione singoli frame 352x240
 - Nota: Il bit-rate è il numero di bit al secondo che devono essere elaborati per riprodurre il filmato
- MPEG2 è stato sviluppato per comprimere video di qualità televisiva
 - Bit rate: 2Mbps o superiori
 - Usato per DVD, cable TV, etc.
- MPEG4 è un formato basato sulla compressione degli oggetti presenti nel filmato ed è un formato che permette di gestire più stream (flussi) video
 - Supporta bit-rate variabili, da <64 kbps fino a 4Mbps e più (nel futuro)

MPEG-1 [1/2]

- ◆ **La compressione è basata sia su una codifica dei singoli frame (intra-frame encoding) che di una codifica che coinvolge più frame (inter-frame encoding).**
- ◆ **Intra-frame encoding**
 - Ogni frame viene compresso separatamente
 - Utilizza una tecnica di compressione denominata DCT (Discrete Cosine Transform)
- ◆ **Inter-frame encoding**
 - Sfrutta la similarità tra frame consecutivi
 - Utilizza una codifica predittiva
 - Un dato frame viene rappresentato attraverso una trasformazione di un frame precedente
 - Interpolative coding
 - Utilizza sia I frame precedenti che quelli successivi come riferimento

MPEG-1 [2/2]

- ◆ **MPEG utilizza tre diversi frame per la codifica**
 - I frames: intra-frame coding
 - **Compressione moderata**
 - **Sono usati come punti d'accesso**
 - P frames: predictive-coded frames
 - **Codificati sulla base dei frame di tipo I e di altri frame di tipo P**
 - B frames: bi-directionally predictive coded
 - **Codificati usando I frame I e P successivi/precedenti**
 - **Compressione elevata**



MPEG-2

	Risoluzione	Frame/s	Bit-rate (Mb/s)	Applicazione
L o W	352x288	30	4	VHS
	720x576	30	15	Studio TV
	1440x1152	60	60	HDTV
	1920x1152	60	80	Produzione film

MPEG-4 [1/4]

- ◆ **Scalabilità del bit-rate in funzione della qualità**
- ◆ **Permette di ottenere una compressione migliore di quella dell'MPEG-1**
- ◆ **Codifica basata sul contenuto**
- ◆ **Fornisce un supporto per uno streaming efficiente**

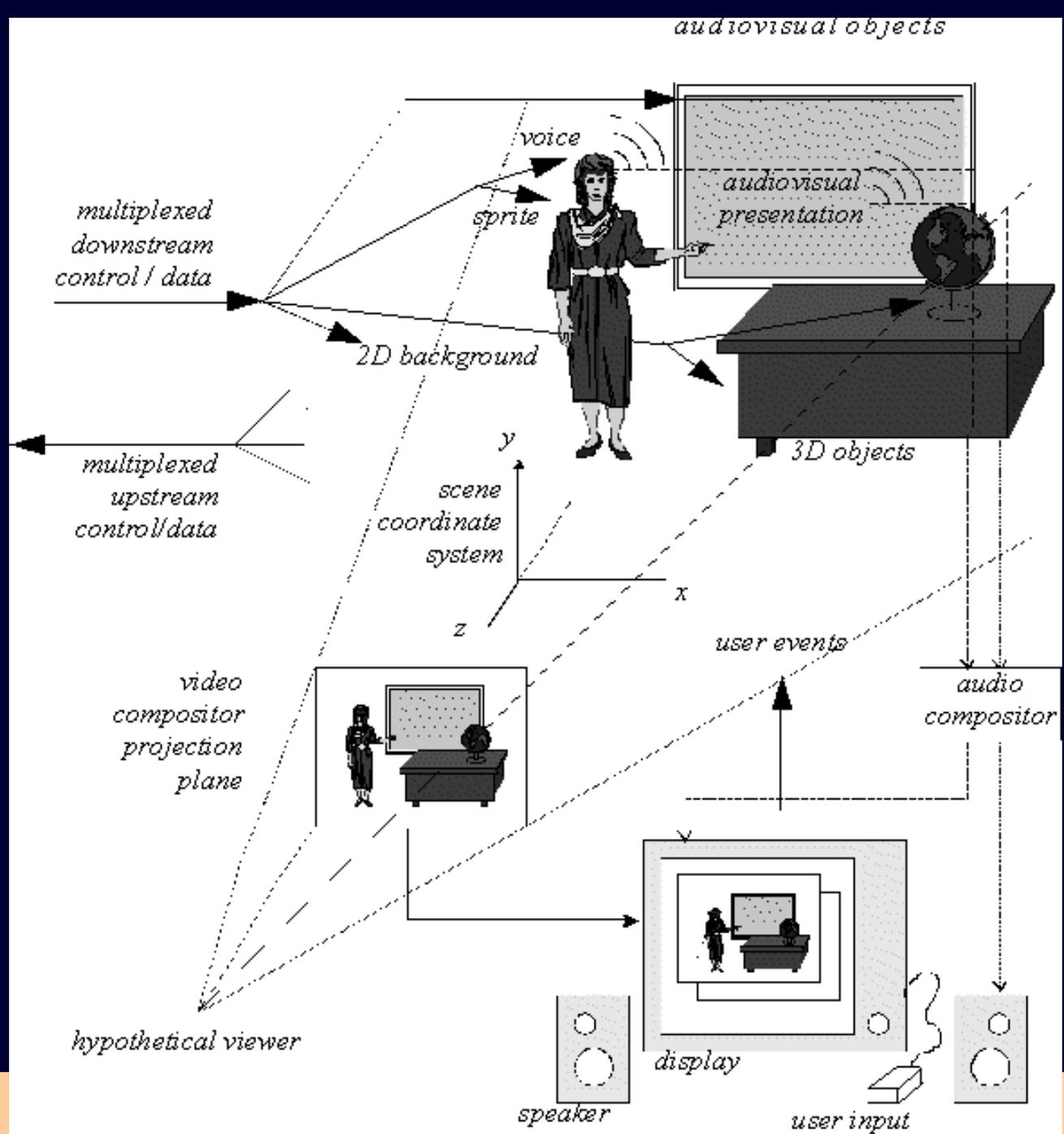
MPEG-4 [2/4]



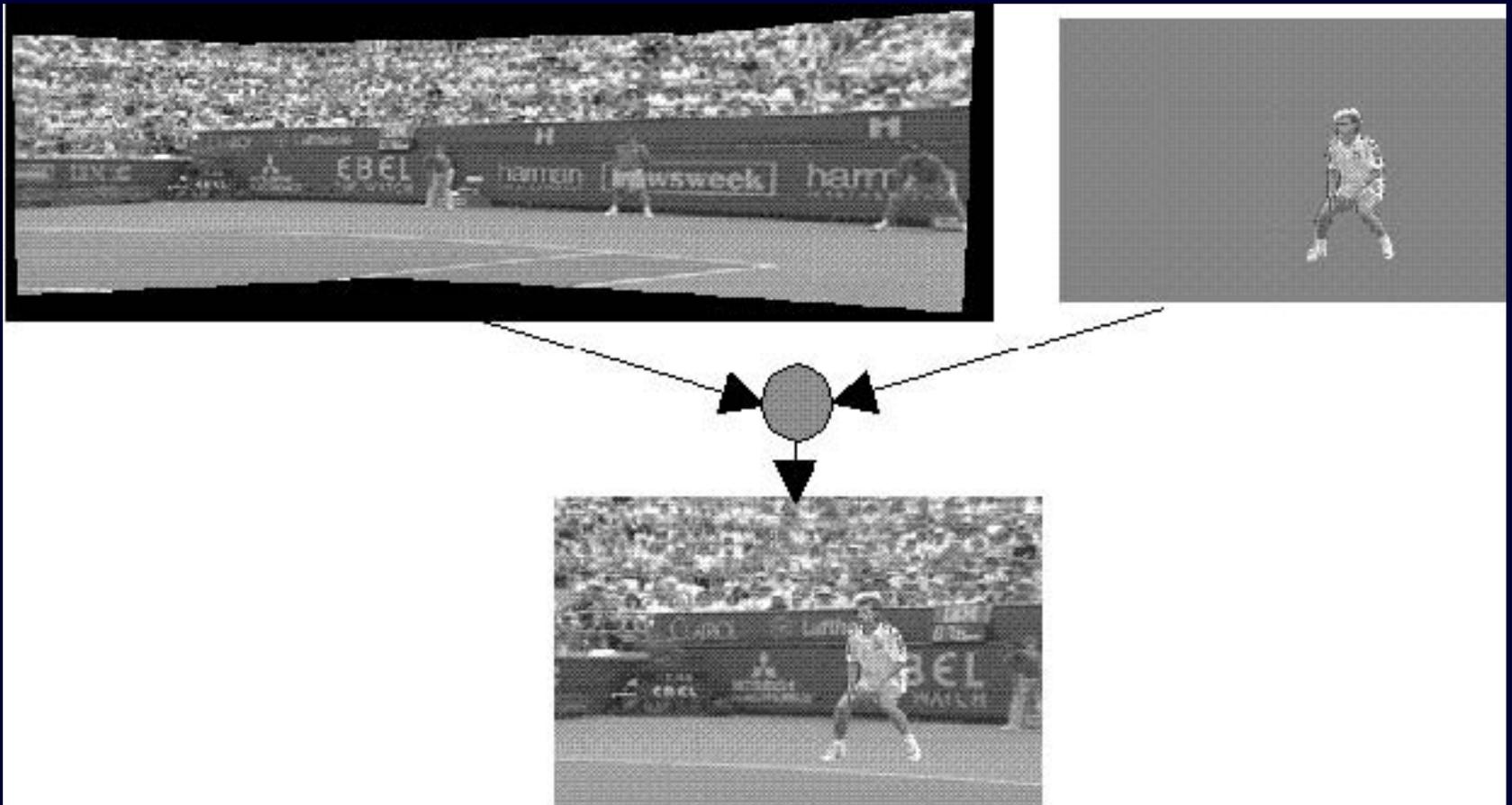
Codifica basata sul contenuto

- Riutilizzo di tecniche di codifica di oggetti
- Adattabilità (codifiche diverse per oggetti differenti)
- Possibilità di codifica ad alta qualità per oggetti o parti significative
- Possibilità di creazione di scene composte
 - **Ad es. Integrazione di contenuto naturale e sintetico**
 - **Tele-presenza**

MPEG-4 [3/4]



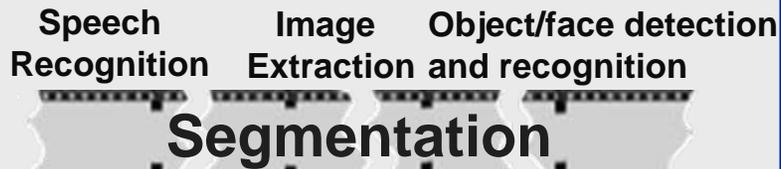
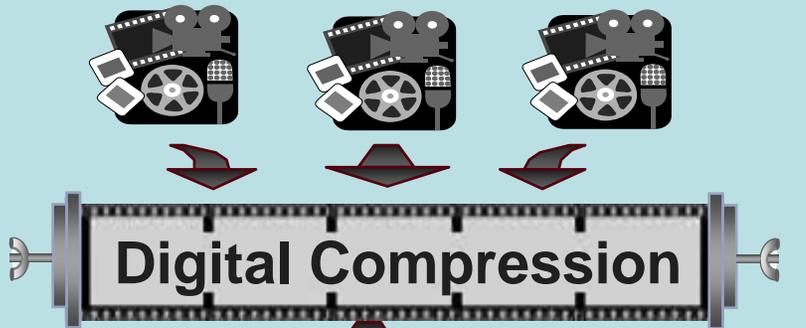
MPEG-4 [4/4]



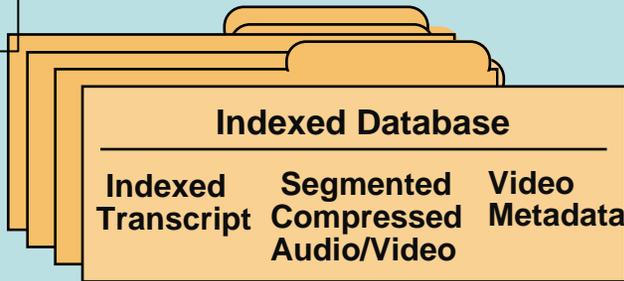
Segmentazione del video

Library Creation

Offline



Insertion of video metadata



DISTRIBUTION TO USERS

Library Exploration

Online

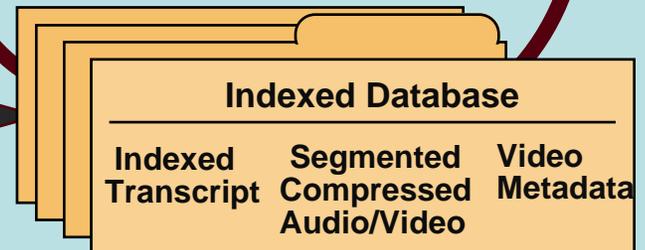


Free text query

Metadata Fields query

Retrieved video segments

Requested Video Segment



Segmentazione del video

- ◆ La segmentazione del video è la decomposizione temporale del contenuto visuale in unità più piccole.
- ◆ I segmenti del video sono normalmente, dal più grande al più piccolo, noti come *sequenze*, *scene*, *shots* e fotogrammi.
- ◆ Lo *shot* è formato da una serie di fotogrammi che identificano una singola azione della telecamera.
- ◆ La *scena* è una sequenza di *shots* che appartengono ad uno stesso contesto, ad esempio la scena di un inseguimento d'auto sempre presente in un film di Hollywood.
- ◆ Infine, una *sequenza* è un segmento video composto da più scene correlate semanticamente, ad esempio un episodio all'interno di un film.

Segmentazione del video (cont.)

- ◆ Il processo di segmentazione si basa sul partizionamento delle sequenze video in shots che sono più facili da identificare.
- ◆ Tipicamente gli shot vengono rilevati automaticamente e vengono rappresentati da uno o più keyframes.
- ◆ Gli algoritmi per l'estrazione degli shots sono utilizzati anche da programmi per l'editing video per facilitare la navigazione nel video e si basano sulla determinazione della transizione temporale da uno shot all'altro.
- ◆ Esistono due tipi di transizioni, quelle improvvise (*cuts*) e quelle per graduali, come la dissolvenza.
- ◆ Un caso particolare di dissolvenza sono la *fade-in* e la *fade-out* che sono delle transizioni del video dal buio o verso il buio rispettivamente.

Identificazione delle transizioni

- ◆ **Le transizioni vengono rilevate analizzando le differenze di due fotogrammi consecutivi.**
- ◆ **Queste differenze possono essere rilevate osservando la variazione di colore dei singoli pixel, se molti pixel dei due fotogrammi sono cambiati allora esiste una buona probabilità di aver individuato una transizione.**
- ◆ **Gli svantaggi di questa tecnica sono due: è molto sensibile ai movimenti della cinepresa, e bisogna lavorare sul video già in forma decompressa. A volte, invece di analizzare singoli pixel, si analizzano alcune statistiche di blocchi di pixel (colore medio, varianza del colore, etc.).**

Identificazione delle transizioni (cont.)

- ◆ **Un altro tipo di approccio analizza le differenze dell'istogramma dei colori di due fotogrammi successivi. Se la differenza tra i due istogrammi supera una certa soglia predefinita, il punto di passaggio tra i due fotogrammi viene marcato come transizione.**
- ◆ **Algoritmi più sofisticati analizzano la variazione su un tempo più lungo (ossia un numero di fotogrammi maggiore di due) in modo da rilevare anche variazioni lente di una scena (dissolvenze).**

Identificazione delle scene

- ◆ **L'identificazione delle scene, vale a dire il rilevamento di transizioni nel contenuto audiovisivo dal punto di vista semantico, invece che fisico (come quello degli shots), è molto più difficile da ottenere ed ancora oggetto di ricerca. Le soluzioni adottate richiedono un alto livello di analisi del contenuto audiovisivo, e si basano su tre tipi di strategie:**
 - l'identificazione di manifestazioni locali (dal punto di vista temporale) basata sulle regole cinematografiche di produzione, che possano far pensare ad una transizione più macroscopica. Ad esempio, effetti di transizione, comparsa/scomparsa di musica dalla colonna sonora.
 - il raggruppamento secondo vincoli temporali: si basa sull'idea che gruppi di contenuti correlati semanticamente tendono ad essere localizzati temporalmente. Quindi solo gli shots che cadono all'interno di una predefinita soglia temporale vengono eventualmente aggregati in un'unica scena.
 - l'uso di modelli a priori sulla base del soggetto: si affidano alla conoscenza a priori del tipo di contenuto: notizie, sport, etc.

Il concetto di Keyframe

- ◆ In realtà realizzare un sistema che permetta di cercare su tutti i fotogrammi di un filmato è poco vantaggioso e dispendioso. Difatti un video realizzato nello standard televisivo europeo PAL è composto da 25 fotogrammi al secondo.
- ◆ Quindi un'ora di video in PAL è composto da $25 \times 60 \times 60 = 90.000$ fotogrammi, una quantità enorme se si pensa che un sistema di archiviazione può contenere migliaia di ore di video. Inoltre molti di questi fotogrammi sono simili tra loro.
- ◆ Per ovviare a questo problema si utilizza solo un sotto insieme dei fotogrammi del video, i **keyframes** (fotogrammi chiave). Un keyframe è un fotogramma particolare che rappresenta adeguatamente una scena di un video.
- ◆ I keyframes vengono generati durante la fase di identificazione delle scene.

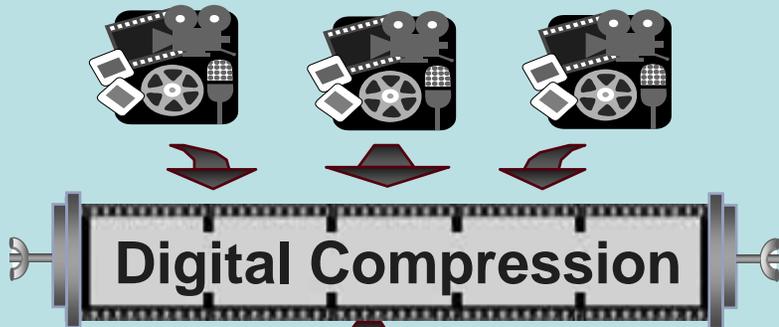
Il concetto di Keyframe (cont)



Indicizzazione automatica e manuale di documenti Audio/Video

Library Creation

Offline



Speech Recognition Image Extraction and recognition Object/face detection

Segmentation

Insertion of video metadata

Indexed Database

Indexed Transcript Segmented Compressed Audio/Video Video Metadata

DISTRIBUTION TO USERS

Library Exploration

Online



Free text query

Metadata Fields query

Retrieved video segments

Requested Video Segment

Indexed Database

Indexed Transcript Segmented Compressed Audio/Video Video Metadata

Indicizzazione

- ◆ **Integrazione tra indicizzazione automatica e manuale**
- ◆ **Indicizzazione automatica**
 - Individuazione keyframes
 - Analisi del movimento
 - Riconoscimento automatico di oggetti
 - Riconoscimento di scritte
 - Riconoscimento di volti
 - Riconoscimento del parlato
- ◆ **Indicizzazione manuale**
 - Aggiunta di attributi alle diverse parti del video (video completo, scene, singoli frame)

Quale è lo scopo dell'incizzazione del video?

- ◆ **Il processo di indicizzazione fornisce una “descrizione” del contenuto video che può essere utilizzata per il processo di ricerca**
- ◆ **Si possono distinguere tre diverse categorie di descrizioni video**
 - Keywords che descrivono il video nel suo complesso
 - Proprietà visuali del video o di sue componenti
 - Informazione semantica

Indicizzazione automatica e manuale

- ◆ **L'obiettivo consiste nel fornire una indicizzazione completamente automatica che sia**
 - Veloce
 - Affidabile (indipendenza dall'utente, pochi errori)
- ◆ **In molti casi questo non è possibile**
 - Complessità del task (ad es. L'interpretazione semantica del contenuto di uno shot può risultare impossibile)
 - L'informazione richiesta non è presente nel video (ad es. Data di creazione, luogo nel quale il video è stato registrato, ecc.)

Information that cannot be extracted automatically

- ◆ **Background information, e.g.**
 - Creation date
 - Author
 - Names of the actors
 - Ecc.
- ◆ **Semantic information**
 - Relations among different shots
 - Interpretation of the meaning of a shot
 - Interpretation of the meaning of a frame
- ◆ **All this type of information must be provided manually, possibly by using a specific tool**

Information that can be extracted automatically

- ◆ **Features that can be extracted from the entire video,**

- e.g. frame rate, resolution, b&w or color video, etc.

- ◆ **Features that are associated to the audio part**

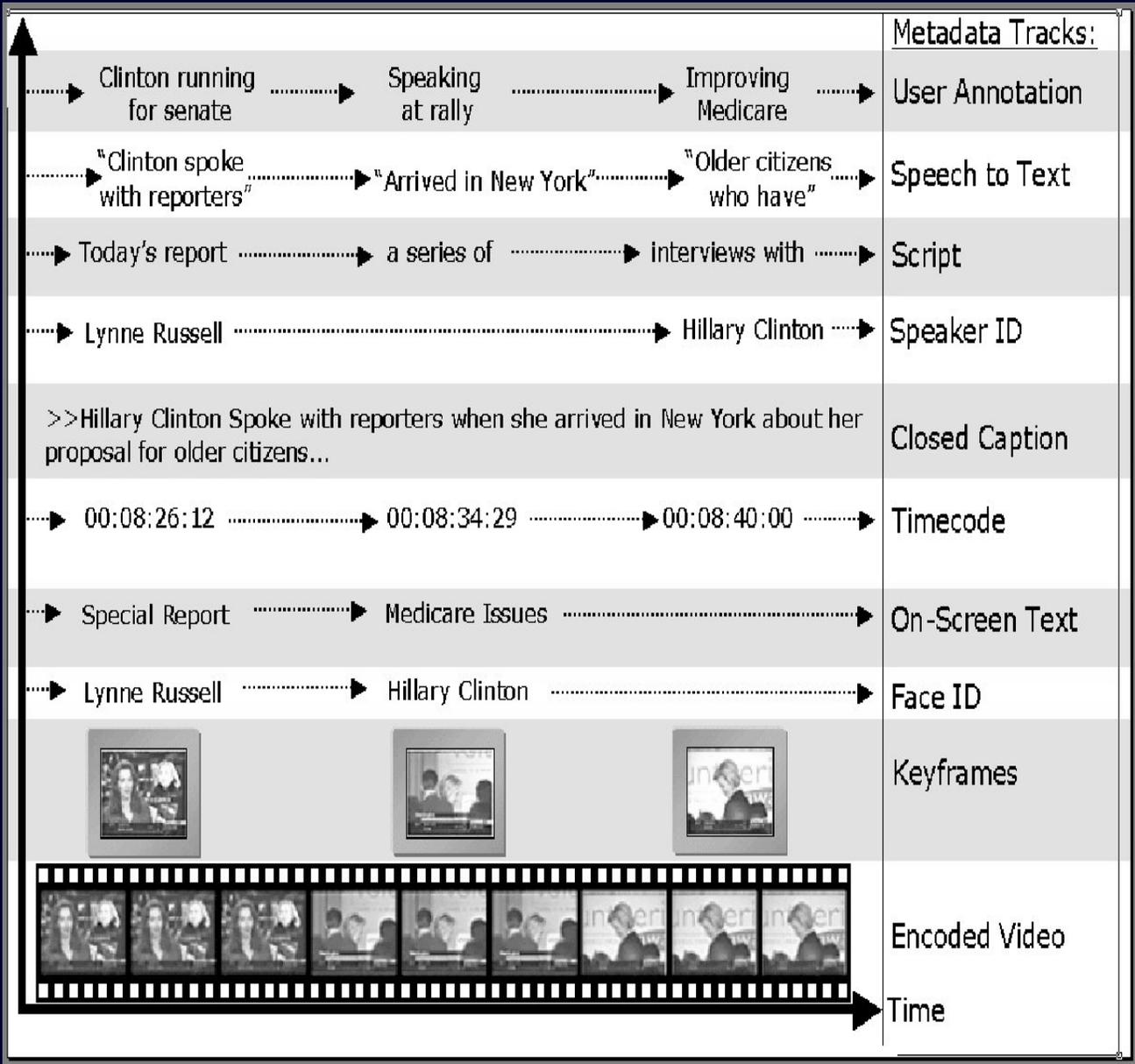
- e.g. the transcript of the speech

- ◆ **Features that can be extracted from each shot**

- e.g. object track, camera movement, recognition of specific objects, recognition of faces, text captions, key frames

- ◆ **Features that can be extracted from each frame**

- these are typical image features, such as color distribution, texture, object's shapes, etc.



Video is indexed by using different tracks, which are time-synchronized

Analisi del movimento

- ◆ **Considerare il video come una semplice sequenza di immagini può essere riduttivo, infatti i fotogrammi di un video sono tra loro correlati temporalmente.**
- ◆ **L'estrazione di features del moto forniscono un modo efficace per effettuare ricerche attraverso la dimensione temporale. Queste features permettono di descrivere in modo sintetico informazioni sul movimento di oggetti nel video o della cinepresa.**
- ◆ **Un esempio tipico è la feature che descrive la quantità di moto presente in una certa scena. Dato che la quantità di moto è semplicemente uno scalare (vale a dire un numero) è possibile utilizzarla per ridurre il numero di fotogrammi su cui effettuare la ricerca usando la similarità per immagini.**
- ◆ **Ad esempio volendo cercare delle scene di calcio da programmi televisivi può essere utile ridurre l'insieme di fotogrammi da cercare a quelli che posseggono una quantità di moto superiore ad una certa soglia. Dopodichè è possibile selezionare le immagini che contengono ad esempio un prato verde.**

Riconoscimento di scritte all'interno di immagini

- ◆ Il riconoscimento di scritte all'interno di un'immagine non è un'operazione complessa. In una prima fase si identificano le regioni che contengono testo, isolando parti dell'immagine che hanno elementi peculiari dei caratteri tipografici.
- ◆ In seguito i segmenti individuati vengono elaborati in modo da accentuare i caratteri rispetto al resto, aumentando ad esempio il contrasto.
- ◆ Infine la parte estratta viene analizzata da un programma OCR (Optical character recognition) che estrae il testo.

Riconoscimento di scritte all'interno di immagini (cont.)



Identificazione

Identificazione e riconoscimento di volti

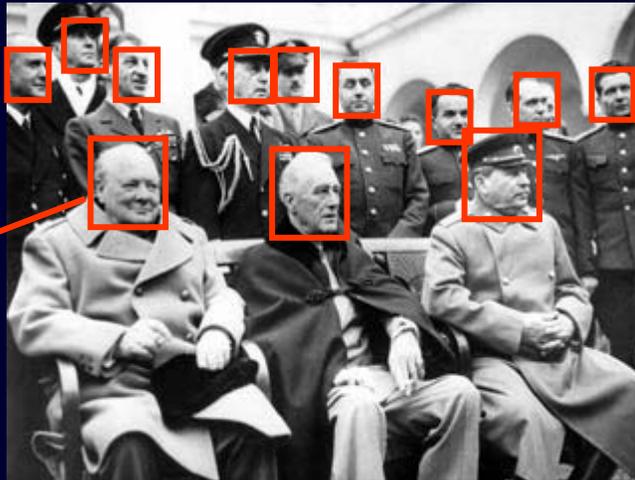
- ◆ **Per ottenere il riconoscimento di un volto è necessario prima indentificarlo.**
- ◆ **Esistono decine di metodi di identificazione di volti, la maggior parte si basano sui colori o sulla ricerca di schemi predefiniti come ad esempio occhi-naso-bocca. I più promettenti utilizzano tecniche che si basano sull'approccio delle reti neurali.**
- ◆ **I problemi da affrontare nell'identificazione di un volto sono svariati. Un volto può essere illuminato in modo insufficiente, trovarsi di profilo, ci possono essere altri volti nella stessa inquadratura o esserci altri elementi di disturbo come animali o cose. Sono quindi frequenti falsi positivi e falsi negativi.**

Identificazione e riconoscimento di volti

- ◆ **Un volto, una volta identificato (tipicamente viene racchiuso in un rettangolo), può essere inviato ad una procedura per il suo riconoscimento.**
- ◆ **Per questa fase è necessario avere uno o più modelli della persona da riconoscere, non è pensabile realizzare un sistema che riesca a riconoscere chiunque. I modelli possono essere utilizzati per creare un base di conoscenza con la quale è possibile annotare un video con i personaggi riconosciuti.**

Identificazione e riconoscimento di volti

winston
churchill



Basi di dati di winston churchill

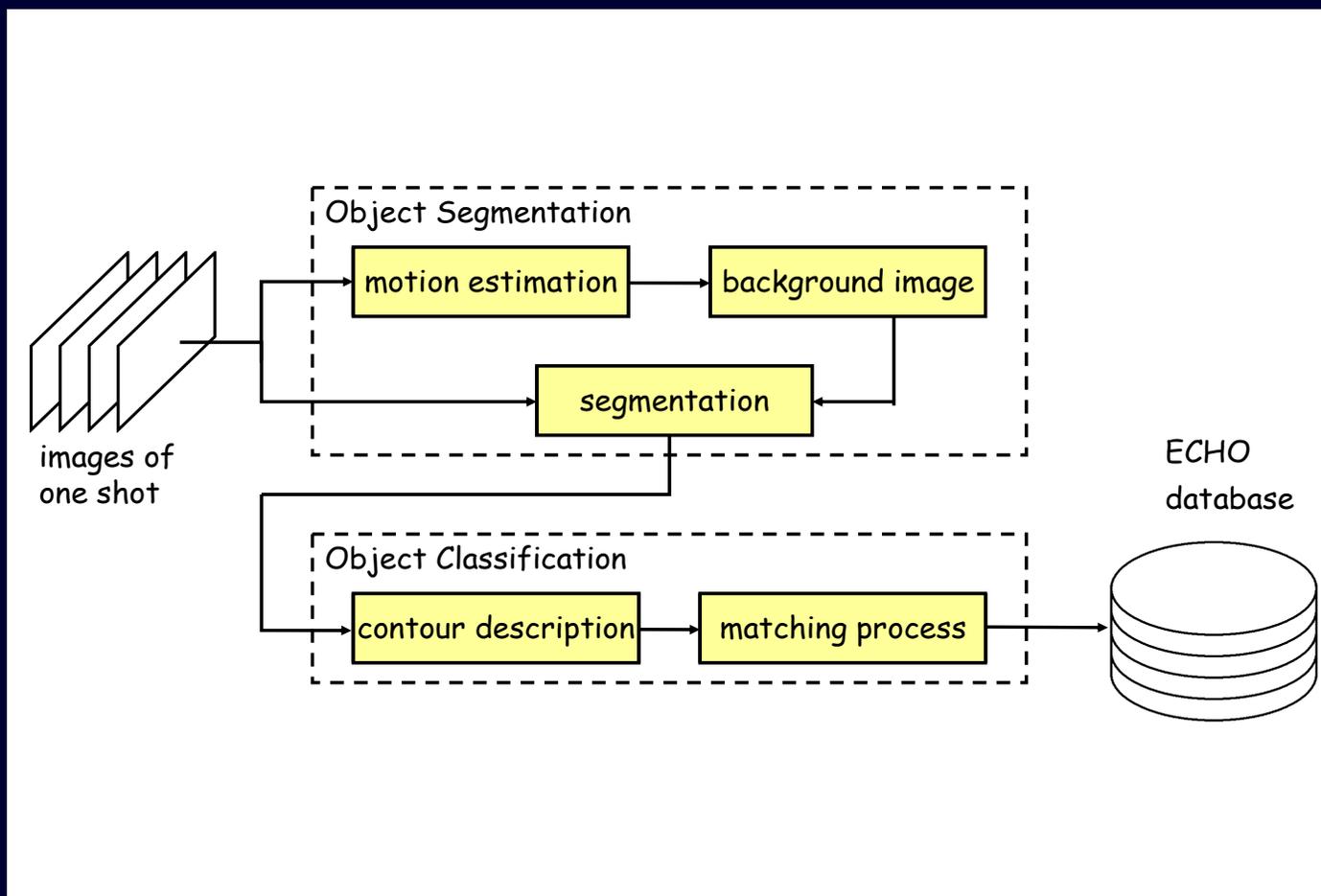
Oggetti

- ◆ **Il riconoscimento del contenuto di un video è senz'altro la sfida più importante che vede impegnati molti ricercatori sia del mondo accademico che industriale.**
- ◆ **Il miglioramento di queste tecniche di riconoscimento potrebbe un giorno colmare quello che è chiamato in inglese *semantic gap*, ossia il divario semantico, che rappresenta oggi il più grosso ostacolo nella ricerca su dati multimediali. Il semantic gap è in pratica la differenza tra quello che l'utente percepisce e quello il sistema automatico riconosce.**

Oggetti (cont)

- ◆ **Oggi si riescono a riconoscere automaticamente abbastanza bene scritte all'interno di un'immagine (e quindi in un video), e si riesce ad identificare un certo insieme di oggetti bene definiti come, automobili, animali, volti, etc.**
- ◆ **Attenzione però a non confondere il concetto di identificazione con riconoscimento. Il primo implica semplicemente l'individuazione all'interno di un'immagine di un'area in cui probabilmente è contenuto un oggetto noto (ad esempio un volto); il secondo, oltre l'identificazione del volto implica anche, appunto, il riconoscimento della persona.**

Phases of Object detection and recognition



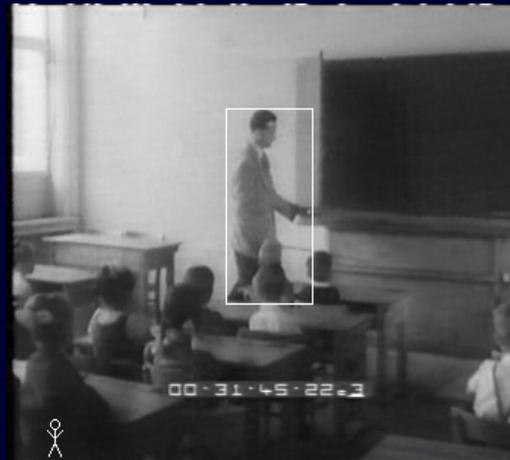
Identificazione di oggetti

- ◆ **L'identificazione di oggetti è una problematica ancora più ampia della precedente.**
- ◆ **In generale nel caso di oggetti si parla semplicemente di riconoscimento piuttosto che identificazione in quanto, quasi sempre, il secondo implica necessariamente il primo. Se devo identificare un cavallo necessariamente lo avrò anche riconosciuto come tale. Questa distinzione è però necessaria per oggetti particolari come le scritte e i volti.**
- ◆ **Quando si cerca di individuare diversi oggetti in un'immagine generalmente si procede ripartendo l'immagine in segmenti significativi che sono poi confrontanti con una base di conoscenza costituita da modelli vari, ad esempio aeroplani, automobili, biciclette, etc.**

Esempio: Automobili



Esempio: persone



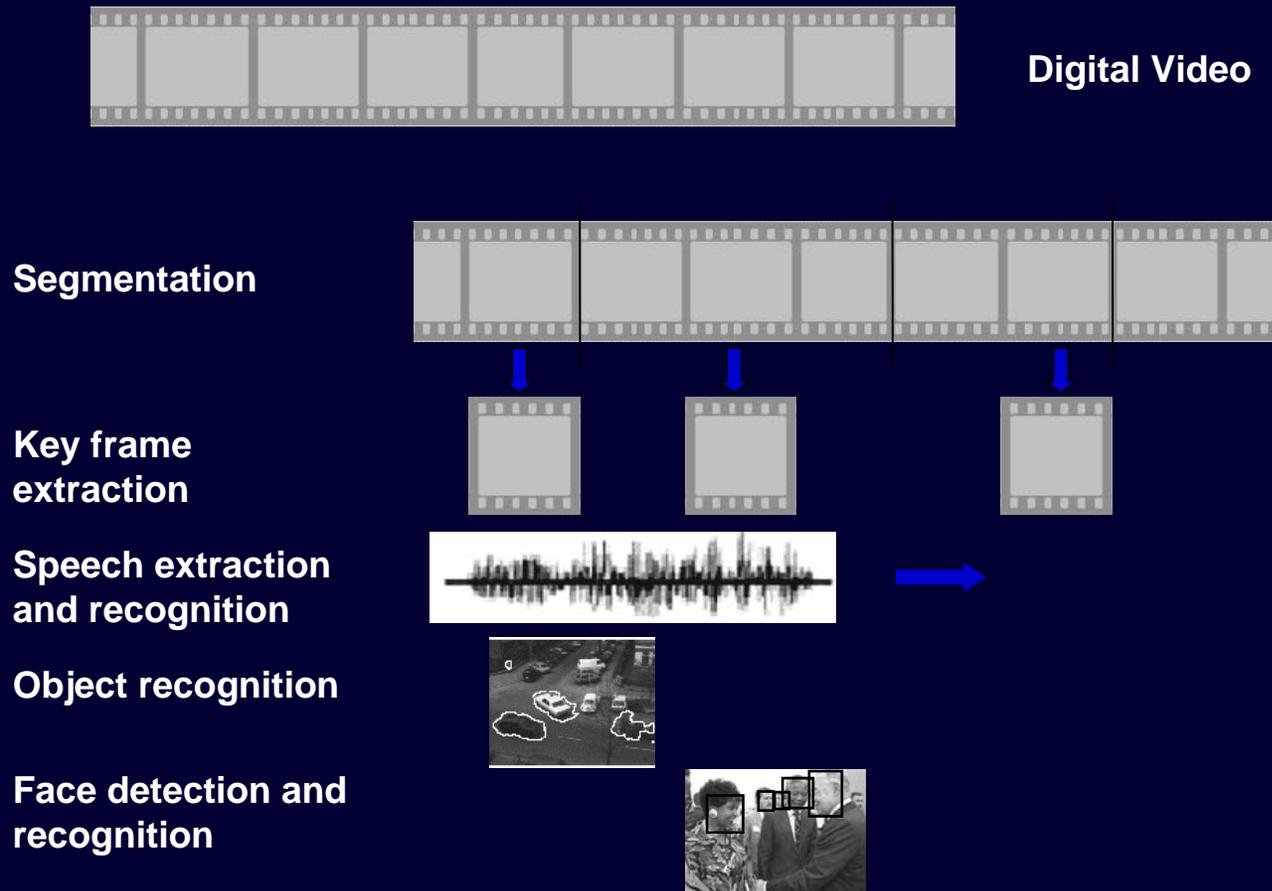
Estrazione del parlato

- ◆ La ricerca di parole o frasi all'interno della trascrizione del parlato può essere sorprendentemente efficace nel recupero di informazioni in documenti audiovisivi. L'utilità del testo estratto dipenderà naturalmente dal tipo di documento audiovisivo.
- ◆ Ad esempio il parlato di un documentario o di un telegiornale è fortemente correlato con quello che è mostrato nel video. Viceversa, in un film il parlato caratterizza poco una scena dal punto di vista semantico, in quanto i personaggi in un film normalmente (come d'altra parte succede nella realtà) non commentano le scene in cui sono coinvolti.
- ◆ Inoltre, mentre in un documentario c'è un solo narratore che parla e che scandisce bene le parole, in un film non succede altrettanto e spesso più persone sono coinvolte nella conversazione e parlano insieme.

Estrazione del parlato (cont.)

- ◆ **L'estrazione del parlato è un processo particolarmente complesso, che si basa su principi simili a quelli utilizzati dai programmi di dettatura.**
- ◆ **A differenza di questi ultimi però l'estrazione del parlato deve fronteggiare diversi problemi come l'ambiente (sovrapposizione di rumori, suoni, musica o riverberi), la qualità di registrazione, le caratteristiche parlatore (sesso, età, proprietà di pronuncia, stato emotivo).**
- ◆ **Ovviamente il risultato di questi algoritmi è affetto da errori, però è stato dimostrato che anche con una percentuale di errori del 50%, l'efficacia della ricerca sul testo è soddisfacente. Il riconoscimento del parlato utilizza le seguenti fasi di riconoscimento:**
 - *Riconoscimento acustico.* Utilizza un modello che descrive il suono dei singoli fonemi del parlato.
 - *Riconoscimento delle parole.* Utilizza un modello del lessico che descrive quali sequenze di fonemi rappresentano parole valide.
 - *Riconoscimento di frasi.* Utilizza un modello linguistico il quale determina la probabilità che una specifica parola sia stata pronunciata, sulla base della lingua del parlatore.

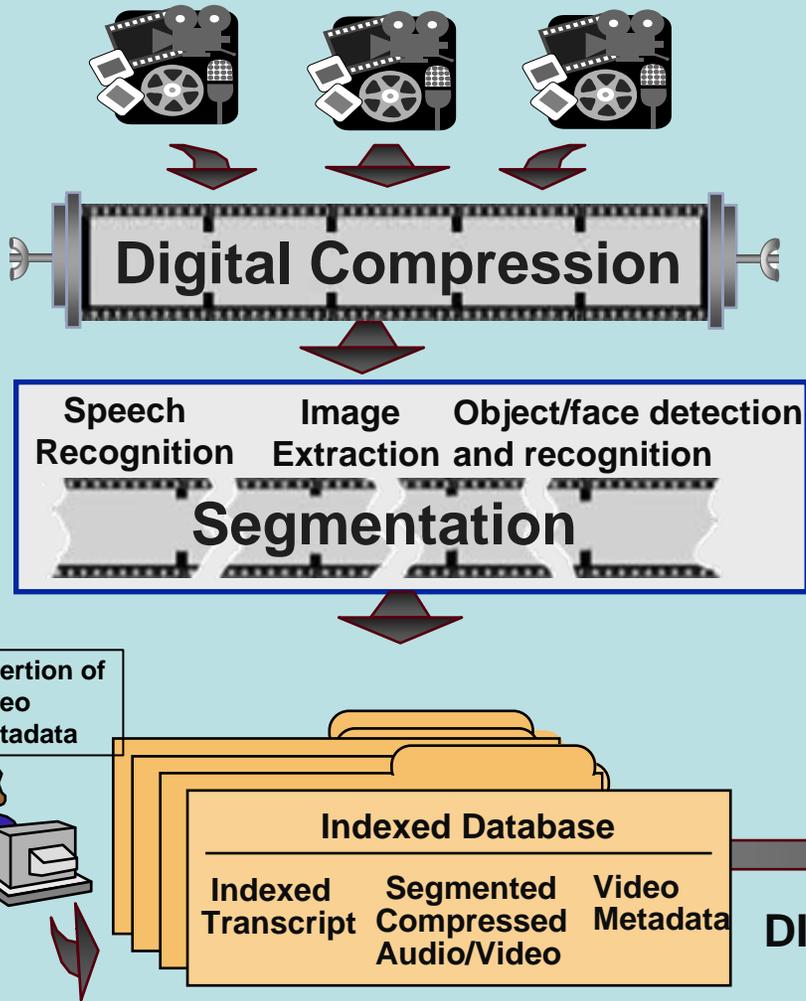
Automatic feature extraction



Retrieval of Audio/Video documents

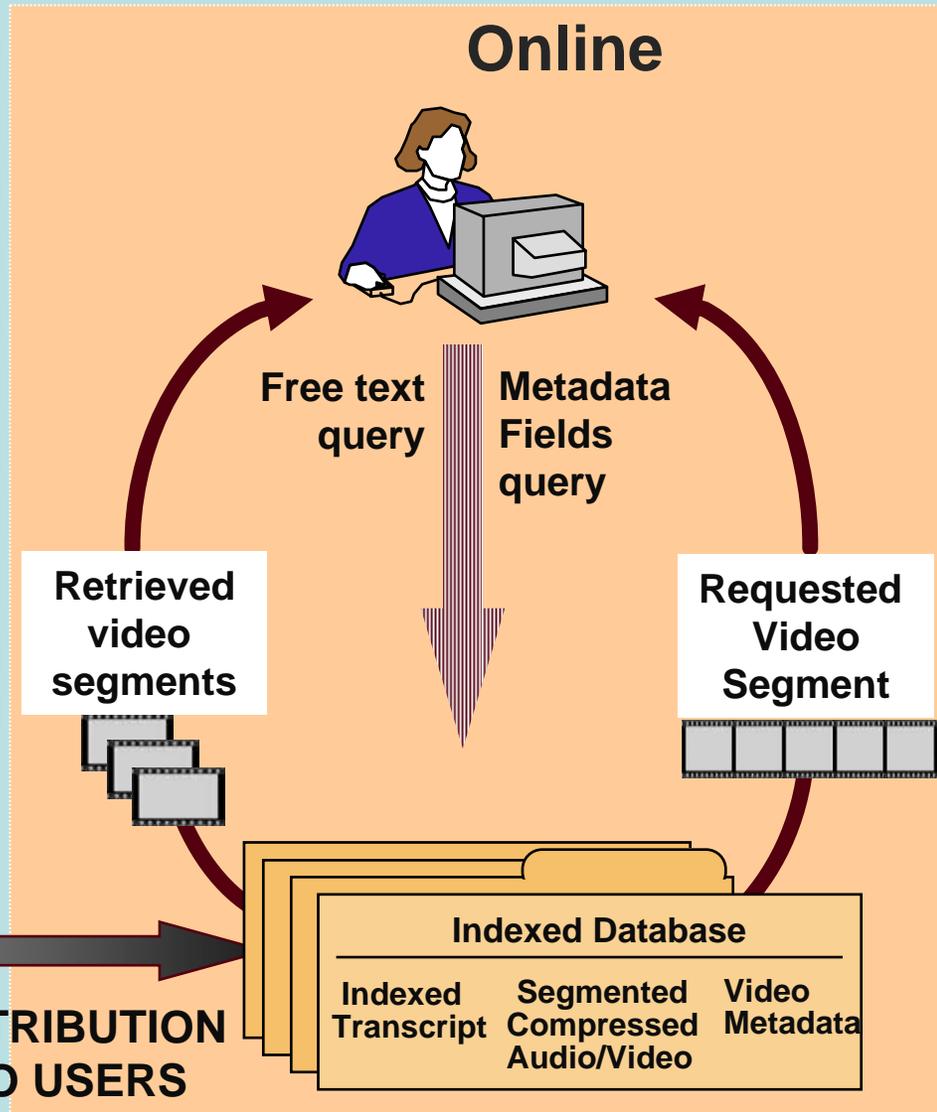
Library Creation

Offline



Library Exploration

Online



Ricerca di video

- ◆ **La ricerca è basata sulla formulazione di interrogazioni che utilizzano le descrizioni dei filmati generate durante l'indicizzazione.**
- ◆ **Sia le descrizioni estratte automaticamente che quelle fornite manualmente possono essere utilizzate.**
- ◆ **Chi formula le interrogazioni normalmente non fa alcuna distinzione tra i due tipi di descrizione.**

Tipi di interrogazioni

◆ Le interrogazioni possono essere espresse su

- Metadati associati all'intero video
 - **E.g.** *find b&w videos produced before II world war by Istituto Luce*
- Metadati associati a shot video
 - **E.g.** *find a shot where the audio transcript contains the words "Attentato Banca Nazionale dell'Agricoltura"*
- Metadati associati a singoli frame
 - **E.g.** *find a video that contains a frame similar to this image [the image is provided as an example]*
- Una combinazione delle precedenti

Informedia – an example

The screenshot displays the CMU Informedia Video Library interface. The main window is titled "CMU Informedia Video Library" and contains a search bar with the query "fires floods earthquakes hurricanes tornadoes". Below the search bar, there are buttons for "Clear All" and "History". The search results section shows "9 of 287 results: any of *fires floods earthquakes hurri tornadoes.*" and a prompt to "Click on a word to highlight it in yellow." Navigation buttons include "Prev. Page", "Next Page", "Go to Page...", and "Visualize All...".

The "Search Results (Page 1 of 32)" section displays a grid of video thumbnails. The "Visualization of search results set containing 287 documents" window is active, showing a scatter plot of documents. The plot is labeled with "fires", "floods", "hur...", "earth...", and "tom...". The visualization includes a legend with checkboxes for "fires", "floods", "earthquakes", "hurricanes", and "tornadoes". The legend is checked for "fires", "floods", "earthquakes", and "tornadoes".

The visualization window also includes a "Display lines toggled. 287 visible documents." message and a "Color Code By:" section with radio buttons for "Minimum Value", "Average Value", and "Maximum Value". The "Relevance (All: 0 - 100)" checkbox is checked. The "Within" field is set to "Date (All: 07/01/99 - 09/30/99)". Other filters include "Size (All: 0:02 - 30:30)", "Map Hits (All: 1 - 125)", and "Topic Hits (All: 1 - 253)".

Informedia – an example

The screenshot displays the CMU Informedia Video Library interface. The main window is titled "CMU Informedia Video Library" and contains several panels:

- Search Panel:** A search bar with the text "fires floods earthquakes hurricanes tornadoes". Below it are buttons for "Clear All" and "History".
- Search Options:** A section with "Search Options" and "Results Options" tabs. It shows "9 of 287 results: any of *fires floods earthquakes hurr tornadoes.*" and a tip: "Click on a word to highlight it in yellow." Navigation buttons include "Prev. Page", "Next Page", "Go to Page...", and "Visualize All..".
- Search Results (Page 1 of 32):** A grid of video thumbnails, each with a "****" rating.
- Visualization Panel:** Titled "Visualization of search results set containing 287 documents". It features tabs for "VIBE", "Timeline", "Map", and "Topics". The "Map" tab is active, showing a world map with a tooltip for "TURKEY: 46/48 hits active". A legend below the map indicates "Documents with no map associations.".
- Filtering Panel:** Includes tabs for "Visible", "Invisible", and "Inactive". A list of locations is shown: "CHINA", "MEXICO", "TANVAN", and "TURKEY". A "Deactivate selected locations" button is present.
- Advanced Filtering Panel:** Displays "58 visible documents; 4 of 5 words ignored; 25 of 55 locations ignored." and various filter options:
 - Color Code By: Minimum Value, Average Value, Maximum Value
 - Relevance (All: 0 - 100): Size-Code Color-Code
 - Date (All: 07/01/99 - 09/30/99): Size-Code Color-Code
 - Size (All: 0:02 - 30:30): Size-Code Color-Code
 - Map Hits (All: 1 - 125; Shown: 3.2 to 125): Size-Code Color-Code
 - Topic Hits (All: 1 - 253): Size-Code Color-Code

ECHO Retrieval Interface

The screenshot shows the ECHO Retrieval Web Service interface in Microsoft Internet Explorer. The interface is divided into several sections:

- Search Form (Left):** Includes fields for Label (In Titles, In Content, In Dates, Collocation-ID), Value (*, 1933), and a Submit button. It also has a 'show 500 hits' option and a 'sorted by' dropdown.
- List of Retrieved Items (Bottom Left):** A list of items with their titles, durations, and dates. The first item is 'Francoforte sul Meno (Germania). I primi lavori stradali della progettata rete tedesca'.
- Detailed View (Right):** A detailed view of a selected item, showing its title, series title, series number, genre, abstract, themes, description language, producer name, production date, producer nationality, kind, silent status, color, audio language, collocation, and provider.

Red circles and arrows highlight specific features:

- Red circles around the search form, the list of items, and the detailed view.
- Red arrows pointing to the search form, the list of items, and the detailed view.
- Red circles around the 'Metadata Edit' and 'Transcript' icons in the top navigation bar.
- Red circles around the 'ECHO' logo and the 'Default' dropdown menu.
- Red circles around the 'Submit' button and the 'show 500 hits' option.
- Red circles around the 'Genre' and 'Newsreel' fields in the detailed view.
- Red circles around the 'Production Date' and 'Production Name' fields in the detailed view.
- Red circles around the 'Audio Language' and 'DE' fields in the detailed view.
- Red circles around the 'Collocation' and 'B024203' fields in the detailed view.
- Red circles around the 'Provider' and 'IL' fields in the detailed view.
- Red circles around the 'Storage ID' and 'B024203' fields in the detailed view.

Integration of Clients and Services in GUI

Views on the material

Attribute Search Retrieval interface

Search for category level of material (Work, Expression, Manifestation, Item)

Detailed view on an item corresponding to the full ECHO data model (structures, links, ...)

List of retrieved items

ECHO Prototype 3 Web Service - Microsoft Internet Explorer

Adresse: http://cms289.8080/MediaArchive/servlet/startView

Kind of Search:
 Nearest Neighbours Range
 Strategy: distance
 Tune Parameter: 0.01
 Hits: 10

Submit
 hits: 10

	DMGGuid: 444C4755494400DF3A88C2F7B10F009518EE8100 Timecode: 00384960 Distance Value: 0.492198
	DMGGuid: 444C4755494400B4499179D81E300054F49CE503 Timecode: 00067280 Distance Value: 0.63827
	DMGGuid: 444C4755494400897DA0F56A9052001782671B03 Timecode: 00061280 Distance Value: 0.638392
	DMGGuid: 444C4755494400897DA0F56A9052001782671B03 Timecode: 00004240 Distance Value: 0.676915
	DMGGuid: 444C47554944000779CB40F7B10F002E6F2A5403 Timecode: 00051960 Distance Value: 0.680515

Work Title: Nozze principesche a Monaco
 Duration: 00:08:20.19
 Date of Ingest: 21.08.2002-14:38:46
 Material: video

ECHO Model Entity

```

graph TD
  AV[AV-Document] --> V[Version]
  V --> M[Media]
  M --> S[Storage]
  V --> Video[Video]
  V --> Audio[Audio]
  V --> Transcript[Transcript]
  Video --> Audio
  Audio --> Transcript
  
```

Title: Nozze principesche a Monaco
 Has Version: [\(To Version object\)](#)
 Has Audio: [\(To Audio object\)](#)
 Has Transcript: [\(To Transcript object\)](#)
 Unit Indication: Whole
 Silent: Sound
 Color: BW
 Person: Kelly, Grace; Ranieri III di Monaco
 Locations: Principato di Monaco; New York; Statua della libert 

Fertig Lokales Intranet

ECHO Prototype 2 Web Service - Microsoft Internet Explorer

Adresse <http://cms289.8080/MediaArchive/servlet/startView>

Video Segmentation

Find * Submit

as of unlimited show 10 hits of type

in archive * sorted by Relevance in ascending order

Hits: 10

- Zuckerfabrik Frauenfeld
- Zirkus Pilatus
- Zirkus Knie: Training im Winterquartier
- Zibelemarit
- Zentrum des schönen Buches in Ascona
- Zelllager im Schnee
- Zeitmessung und Uhrenindustrie
- ZEVENDE NEDERLANDSE KATHOLIEKENDAG
- Yoga

00:06:59.19...00:06:59.21 (00:00:00.02)

00:06:59.22...00:07:02.09 (00:00:02.12)

00:07:02.10...00:07:04.05 (00:00:01.20)

00:07:04.07...00:07:06.06 (00:00:01.24)

00:07:06.07...00:07:07.13 (00:00:01.06)

00:07:07.15...00:07:08.21 (00:00:01.06)

00:07:08.24...00:07:12.10 (00:00:03.11)

00:07:12.13...00:07:15.02 (00:00:02.14)

Fertig Lokales Intranet

ECHO Prototype 3 Web Service - Microsoft Internet Explorer

File Bearbeiten Ansicht Favoriten Extras ?

Zurück Vorwärts Abbrechen Aktualisieren Startseite Suchen Favoriten Verlauf E-Mail Drucken Bearbeiten Diskussion Real.com Links

Adresse http://cms289:8080/MediaArchive/serlet/startView Wechseln zu

Browsing Metadata Editor Export Import Delete

ES CO Video Segmentation

Find: Landwirtschaft wird ersetzt durch Fabrikarbeit, industrielle Landwirtschaft und - weiterverarbeitung Submit

Archive: IL INA Memoriv NAA Themes: 1 - Post War 2 - The World Wars 3 - Sports in the 20th Century 4 - Daily Life 5 - Youth Culture in Europe

show 100 hits Hits:

- me001
- me001
- RAMSI I
- VW Käfer
- Zuckerfabrik Frauenfeld
- Zirkus Pilatus
- Zirkus Knie: Training im Winterquartier
- Zibelemarit
- Zentrum des schönen Buches in Ascona
- Zeltlager im Schnee

Fertig Lokales Intranet

