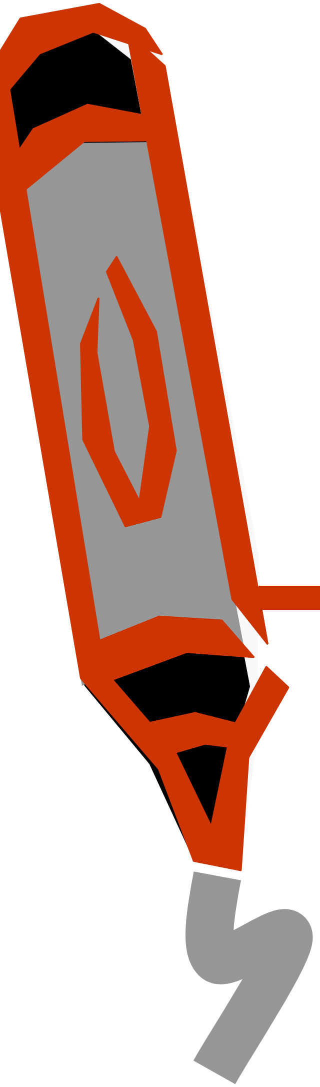


multimedia retrieval

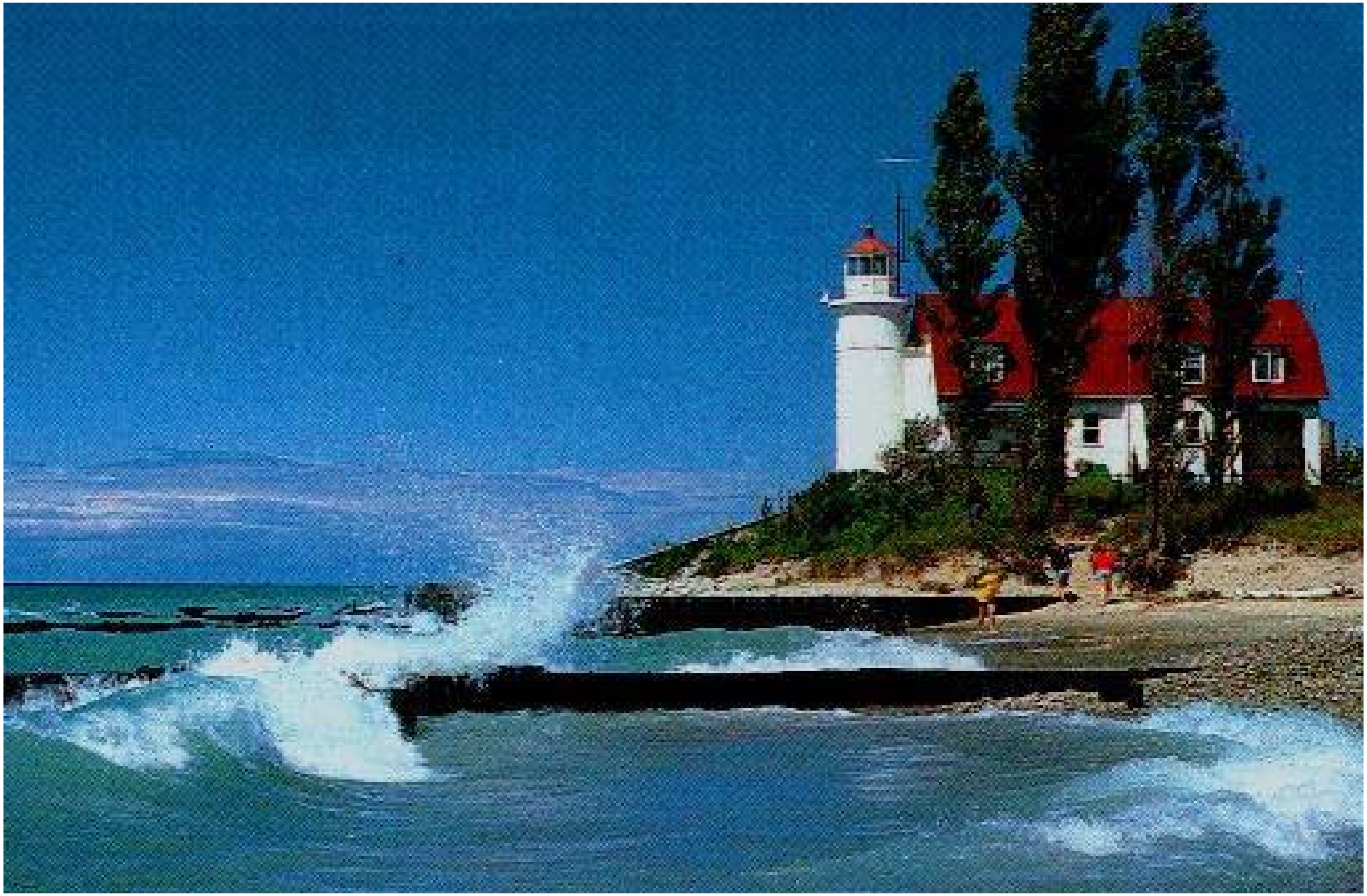
some slides courtesy

Jimmy Lin , University of Maryland
Dr. Ramesh R. Sarukkai, Yahoo! Search

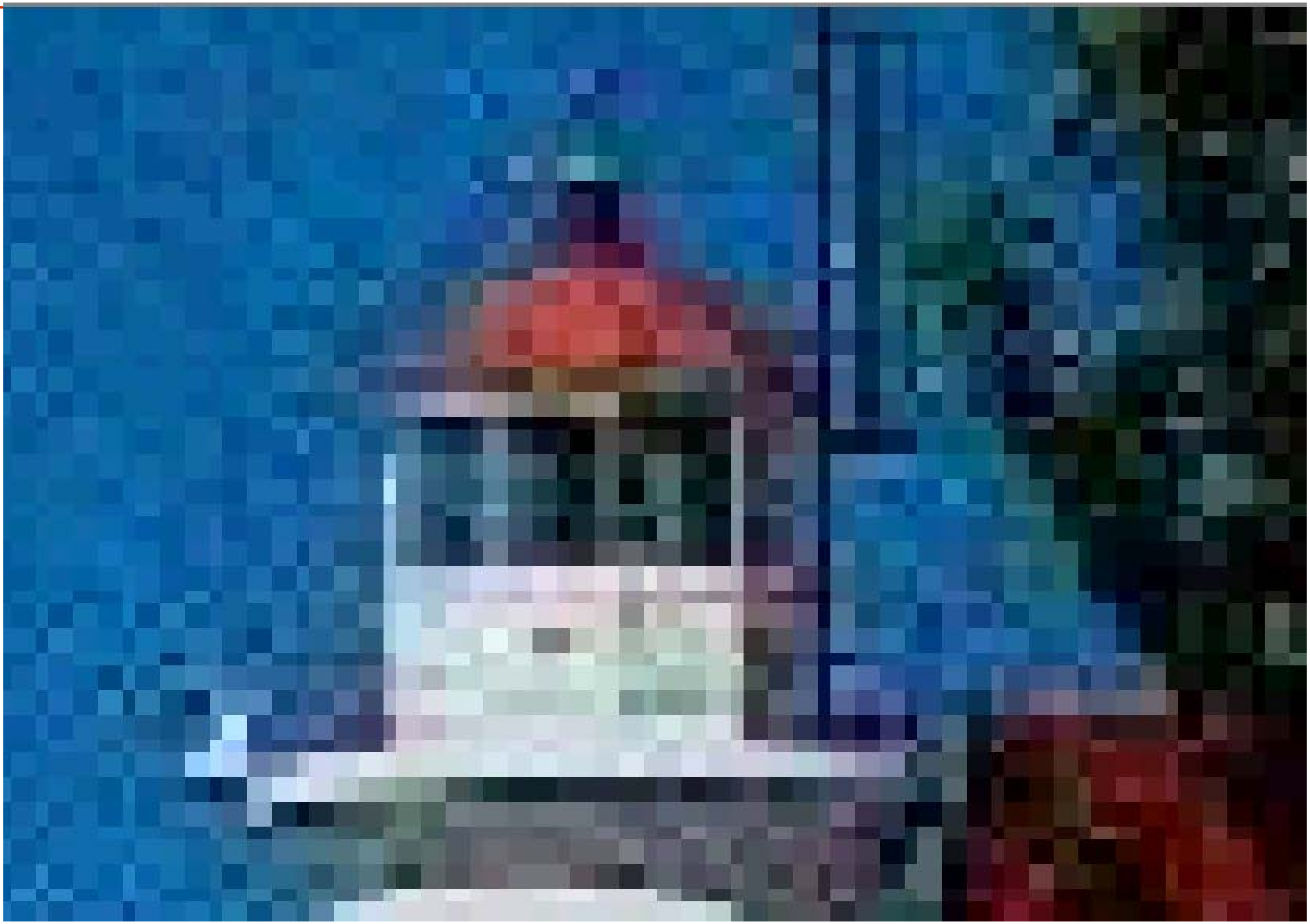
outline

- images
- video
- speech

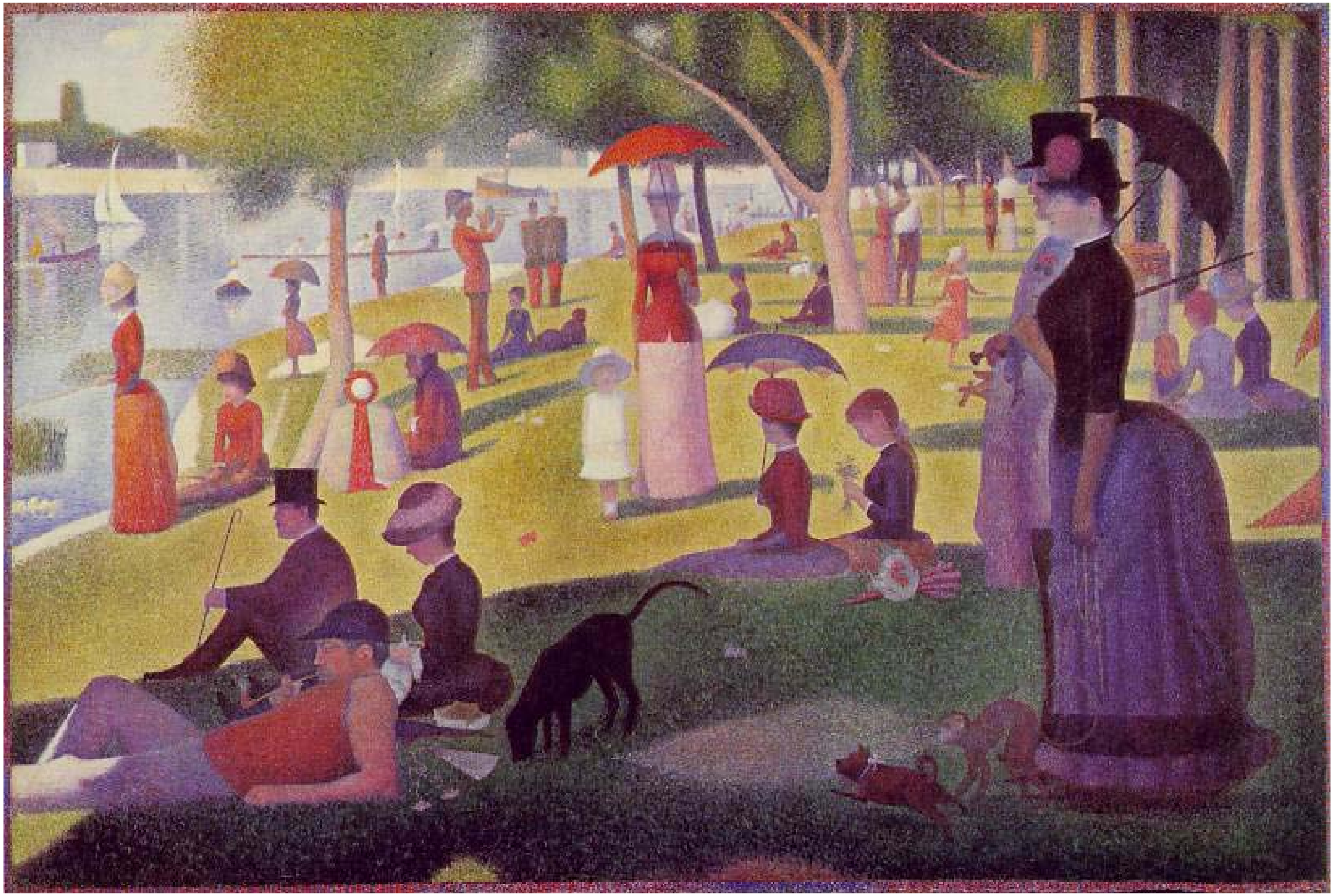
A Picture…



… is comprised of pixels

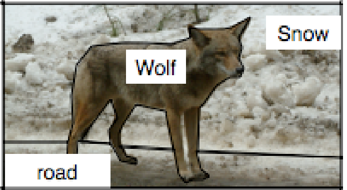
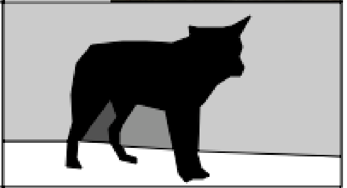


This is nothing new!



Seurat, Georges, A Sunday Afternoon on the Island of La Grande Jatte

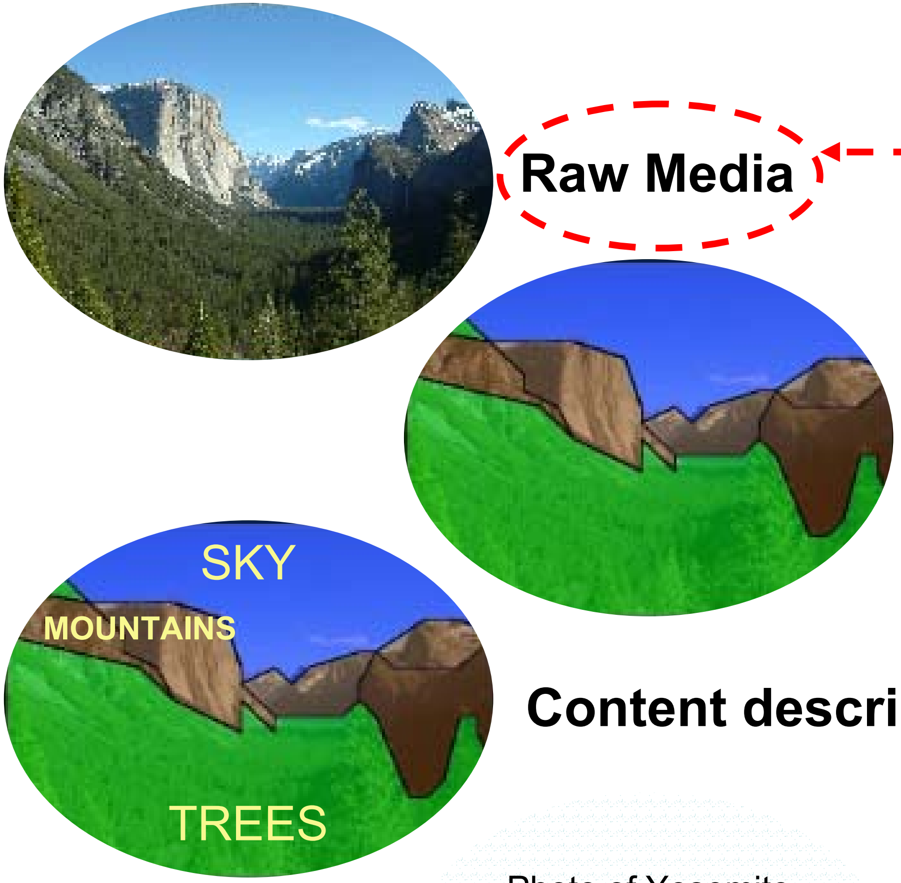
The Semantic Gap

Semantics object relationships and more	Wolf **on** Road **with** Snow **on Roadside in Yosemite National Park, California on 24/1/2004 at 23:19:11GMT**
Object Labels symbolic names of objects	
Objects prototypical combinations of descriptors	
Descriptors feature-vectors	Segmented blobs, Salient regions, Pixel-level histograms, Fourier descriptors, etc...
Raw Media images	

The Semantic Gap

- Content-based retrieval often fails due to the gap between information extractable automatically from the visual data (feature-vectors) and the interpretation a user may have for the same data
  - ...typically between low level features and the image semantics
- The current hot topic in multimedia IR research

The Semantic Gap



Raw Media

This is what we have to work with

Image-level descriptors

Content descriptors

This is what we want

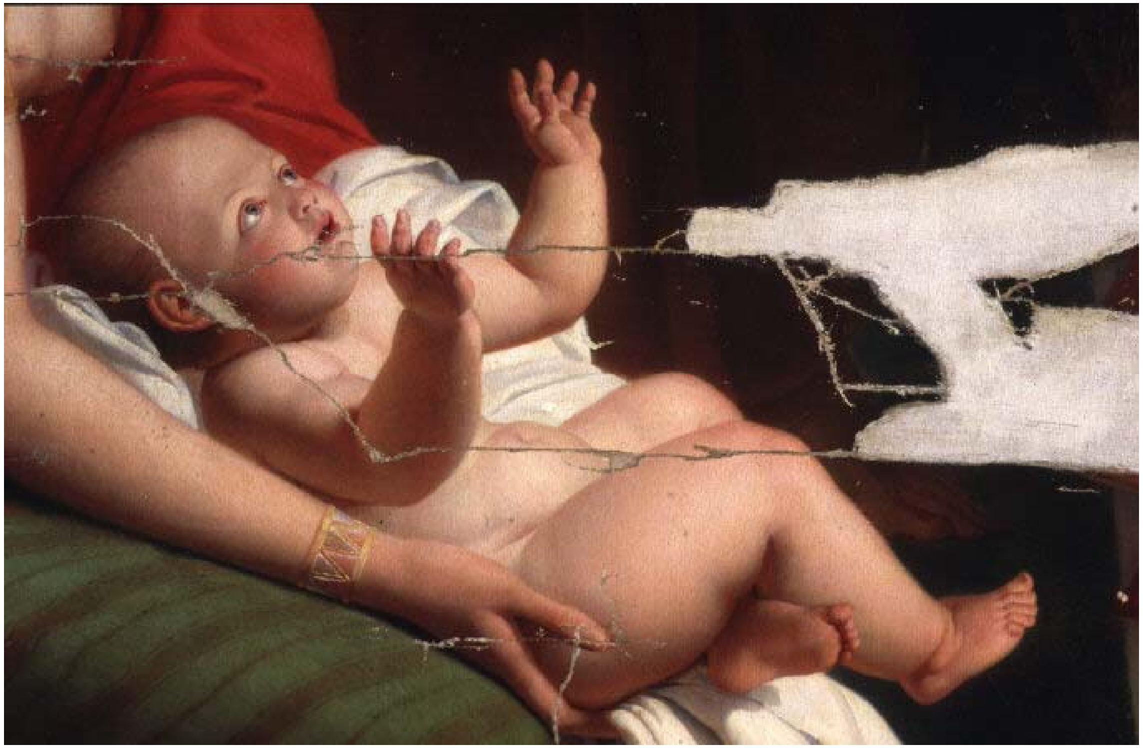
Photo of Yosemite valley showing El Capitan and Glacier Point with the Half Dome in the distance

Semantic content

Sub-image matching

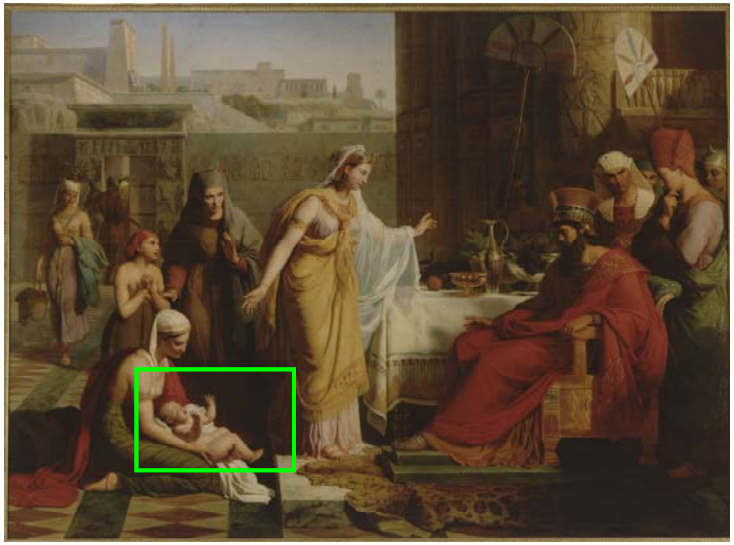
- Given a query image, find the parent image in the database with which it matches, either as a whole or as a part.
- Give location information showing where in the parent the query is positioned
- The images may be very high resolution
- The query and target may be at different resolutions

Example 1: Query



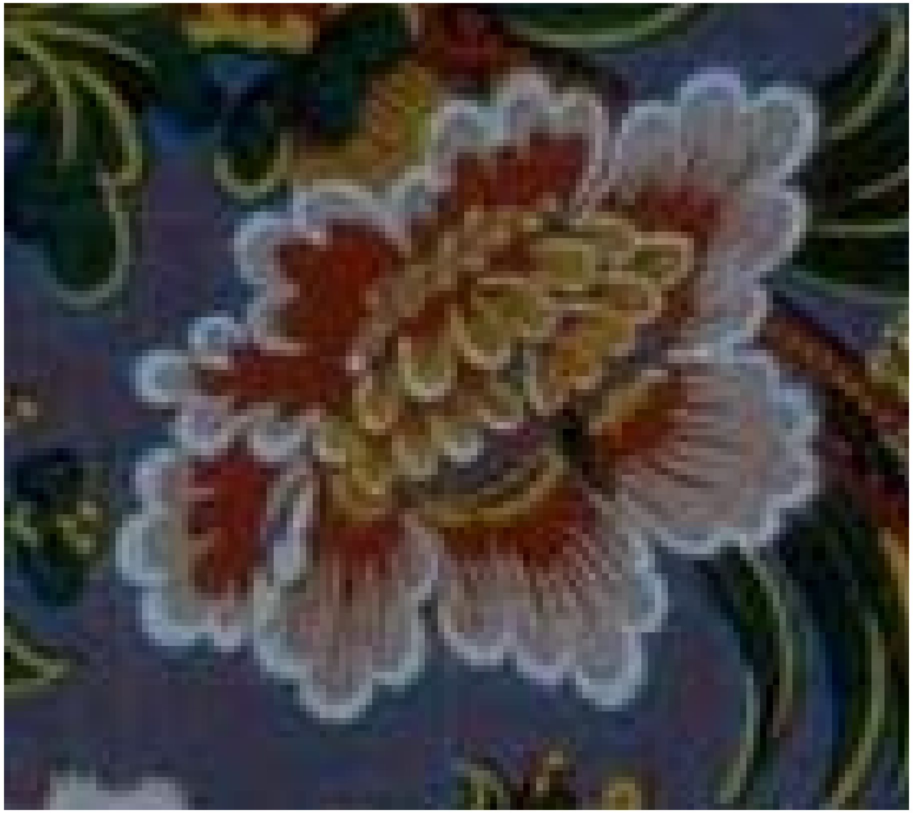
Example 1: Result

- Best matching image with sub-image identified



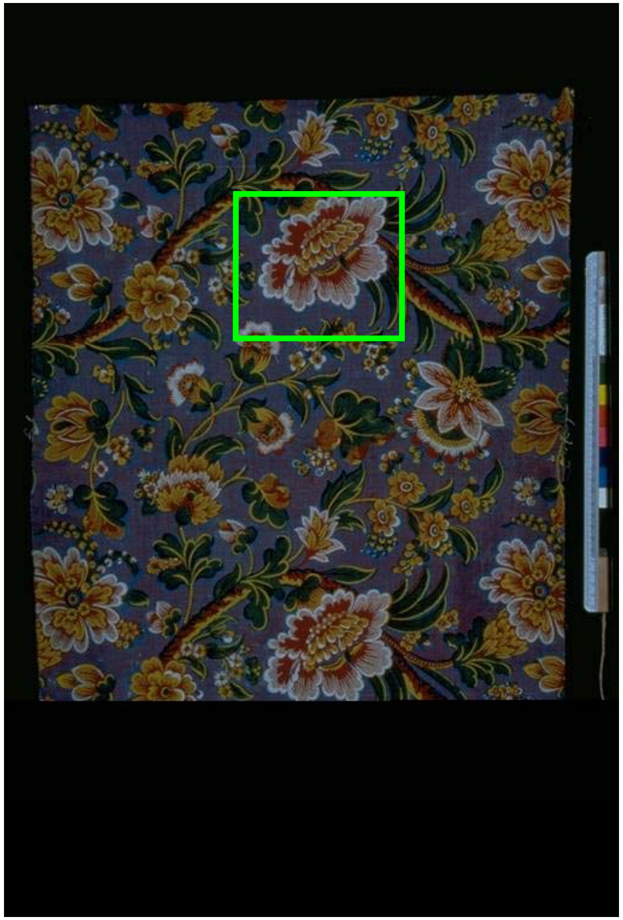
NB. Query is before restoration work, target is a restored image. Query and target image also differ in resolution

Example 2: Query

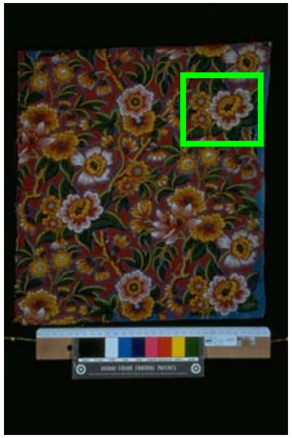
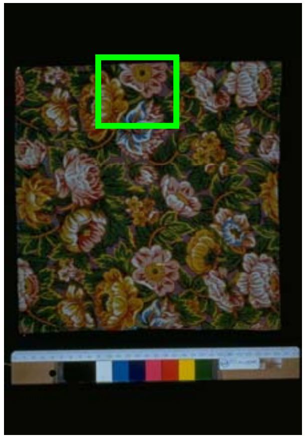


Example 2: Result

Best match found, with sub-image identified

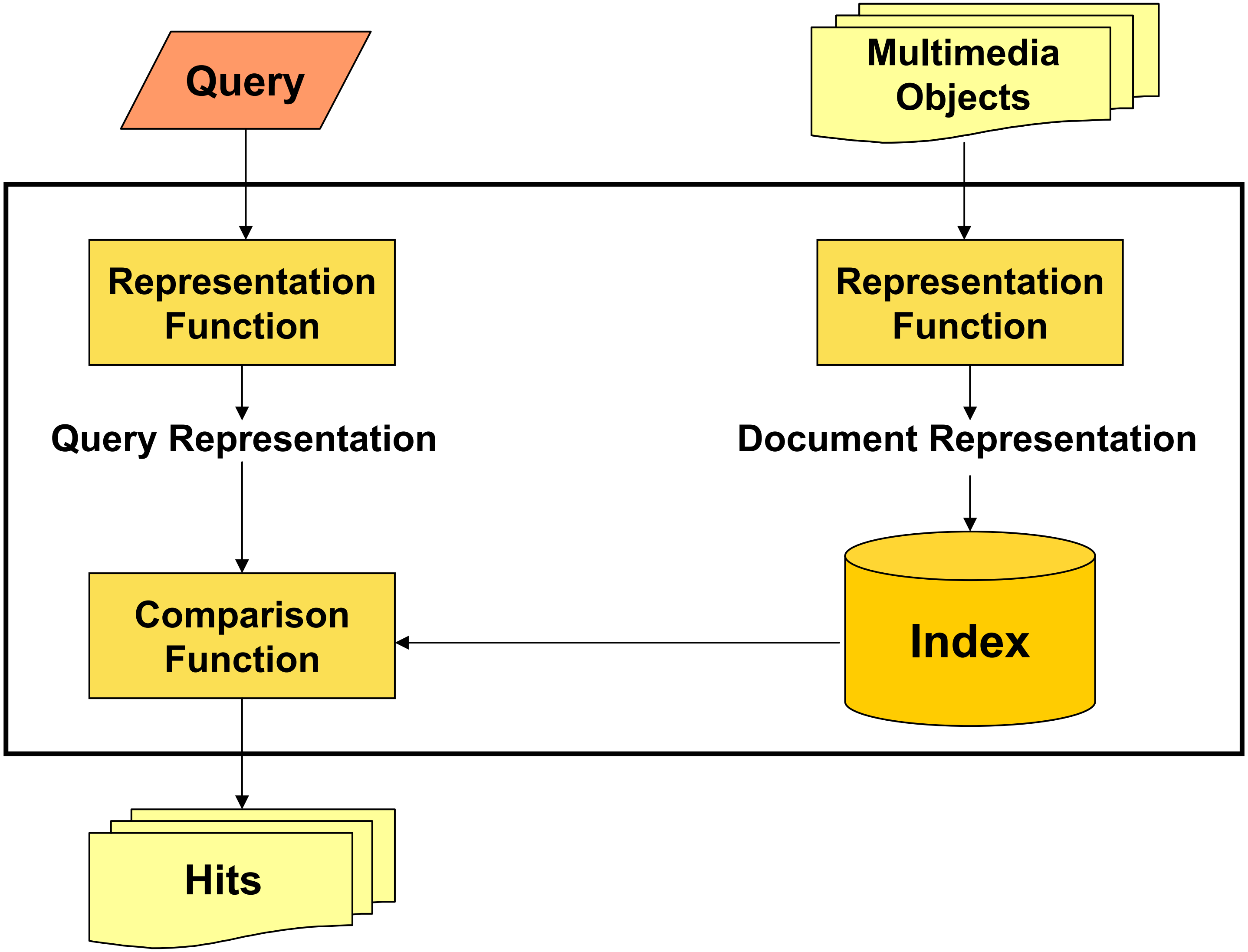


…Subsequent Best Matches

Retrieved results start from top-left to bottom right.

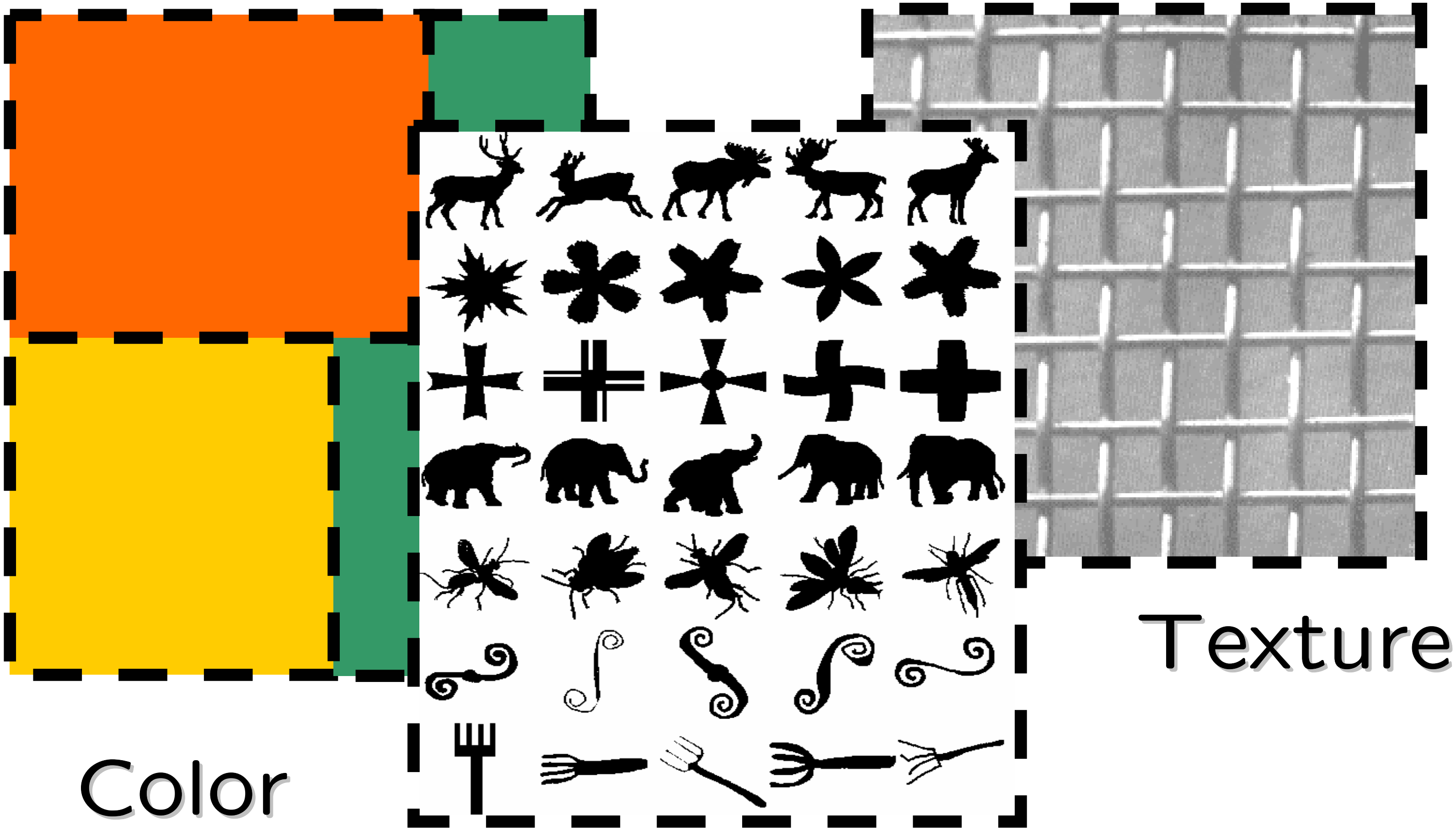
The IR Black Box



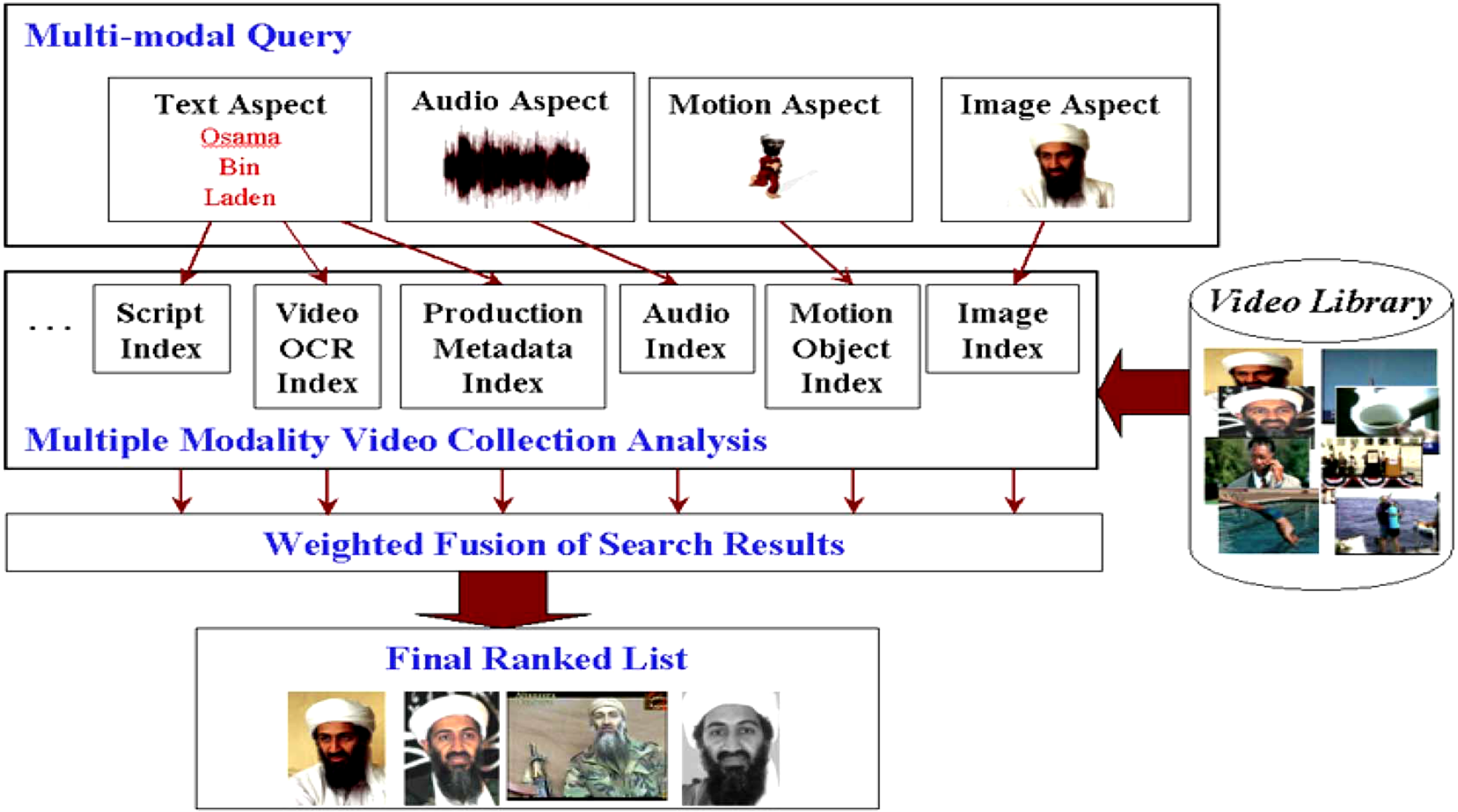
Recipe for Multimedia Retrieval

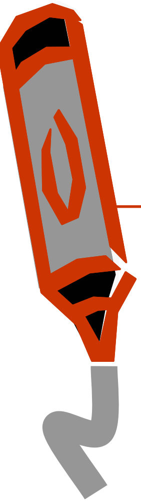
- Extract features
  - – Low-level features: blobs, textures, color histograms
  - – Textual annotations: captions, ASR, video OCR, human labels
- Match features
  - – From “bag of words” to “bag of features”

Visual Features …



Combination of Evidence

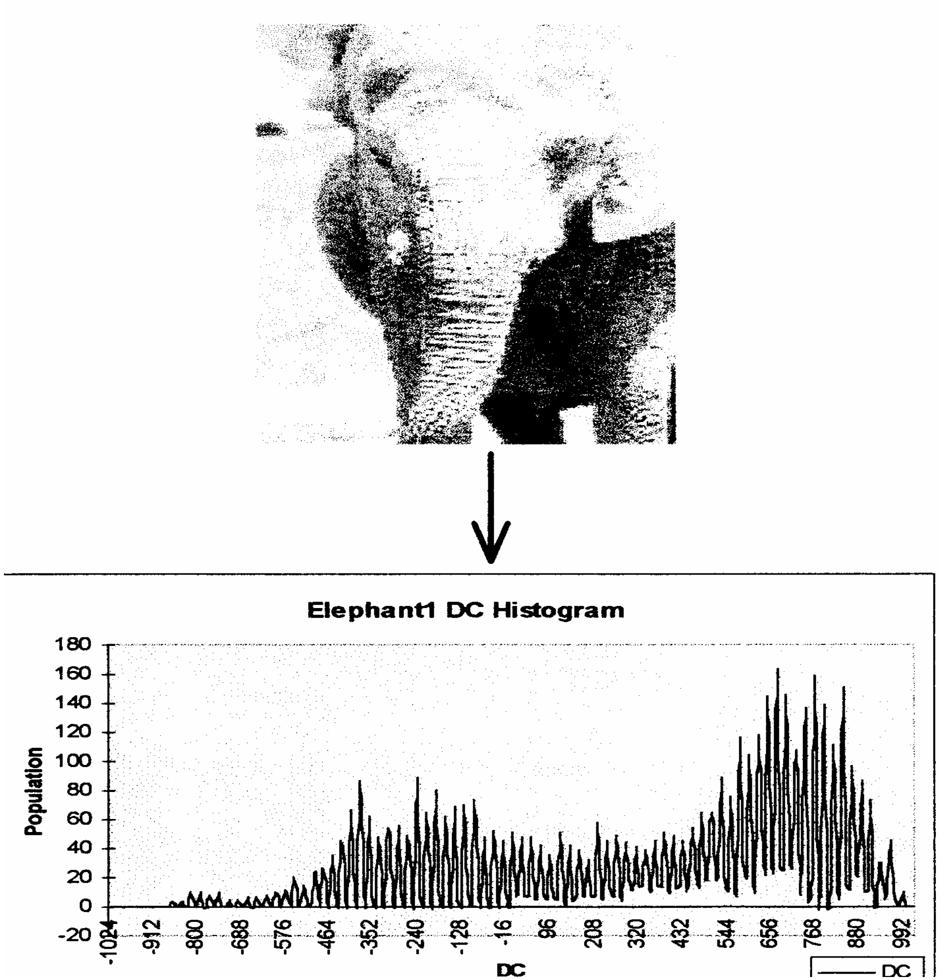




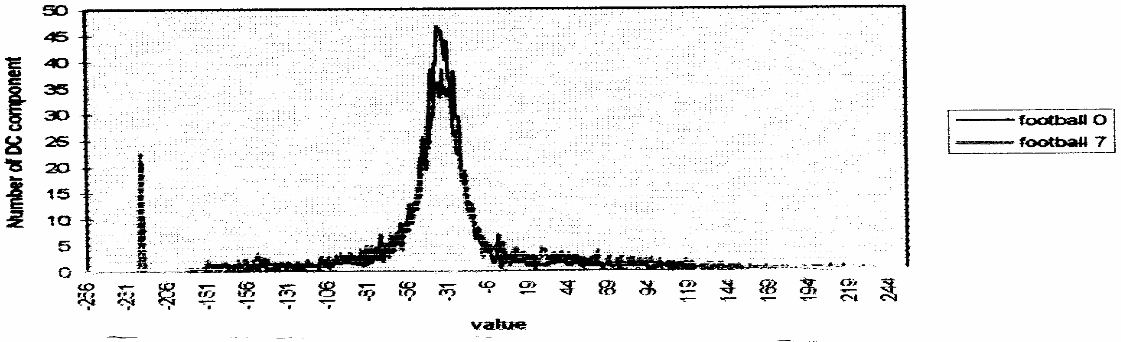
New Algorithm for Similarity-Based Retrieval of Images

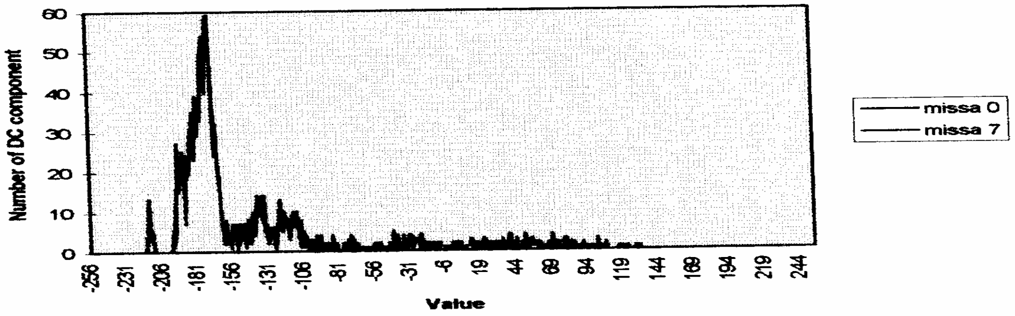
- Images in the database are stored as JPEG-compressed images
- The user submits a request for search-by-similarity by presenting the desired image.
- The algorithm calculates the DC coefficients of this image and creates the histogram of DC coefficients.
- The algorithm compares the DC histogram of the submitted image with the DC histograms of the stored images.

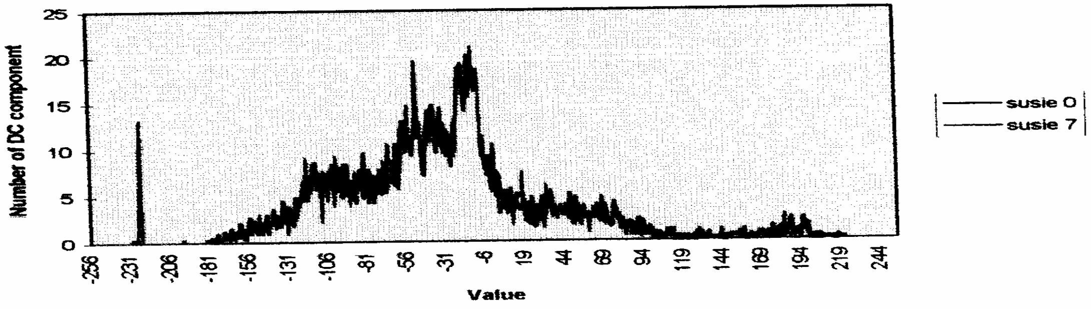
Histogram of DC Coefficients for the Image "Elephant"



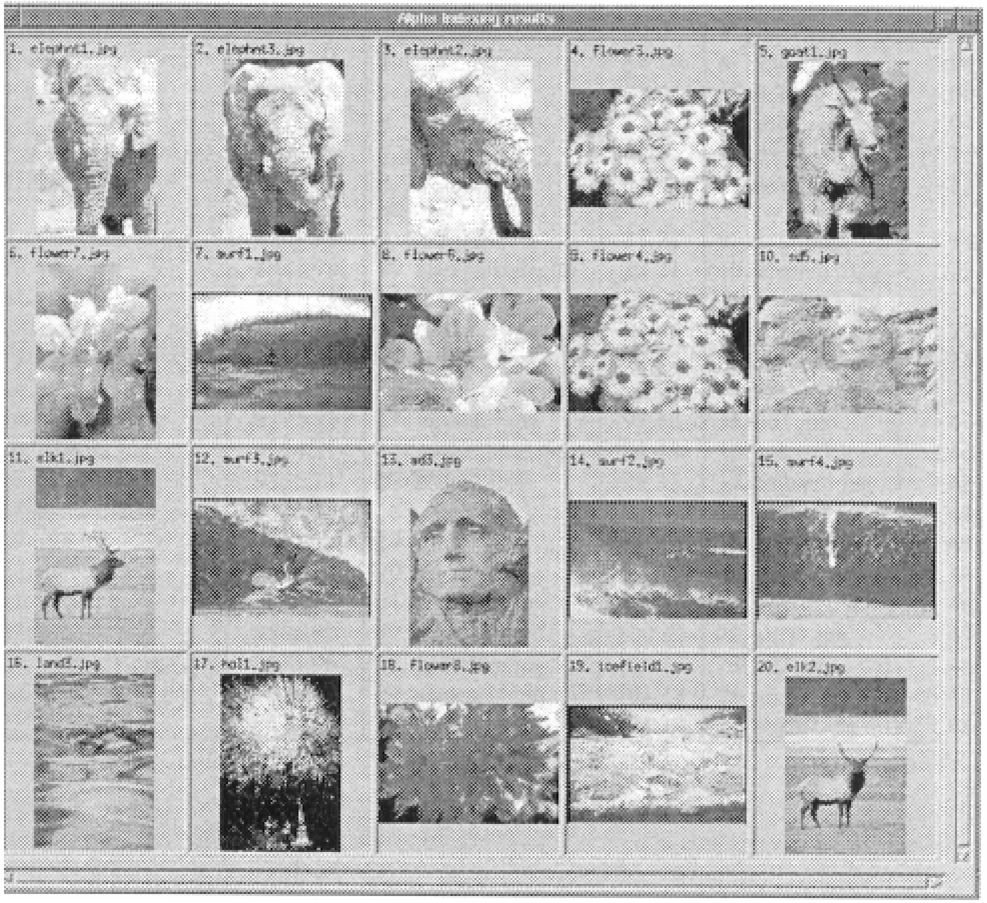
Comparison of Histograms of DC Coefficients







Example of Similarity-Based Retrieval Using the DC Histograms



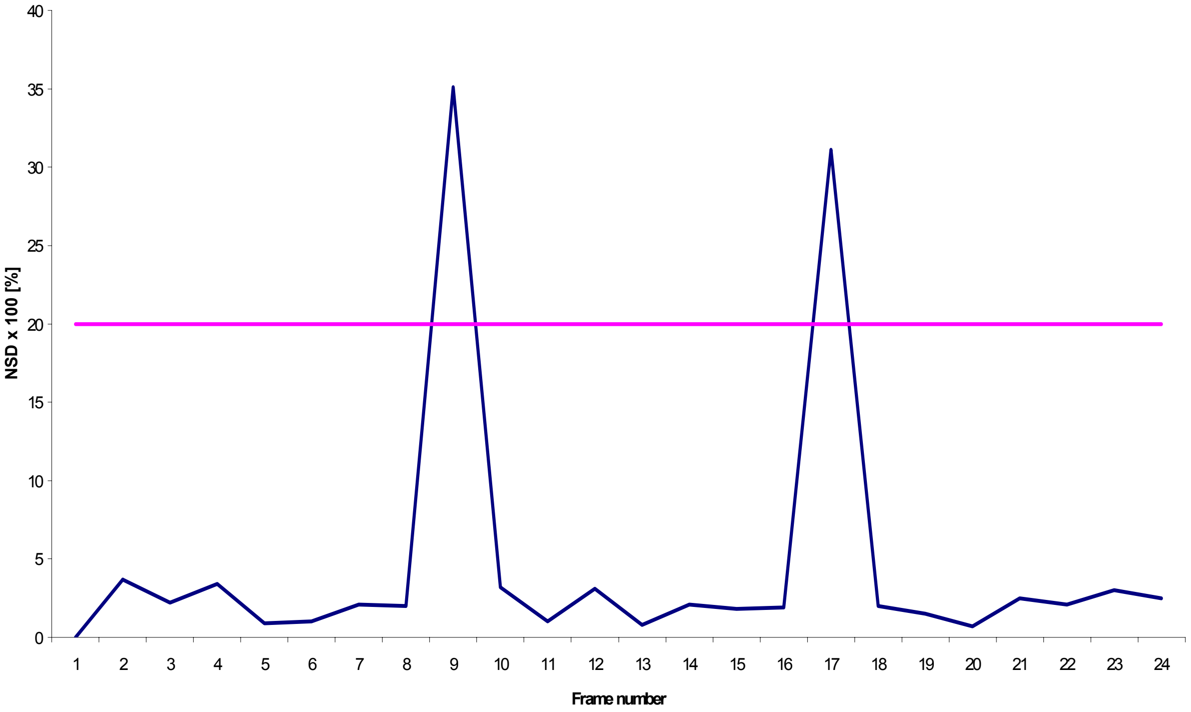
Similarity-Based Retrieval of Compressed Video

- Partitioning video into clips - video segmentation

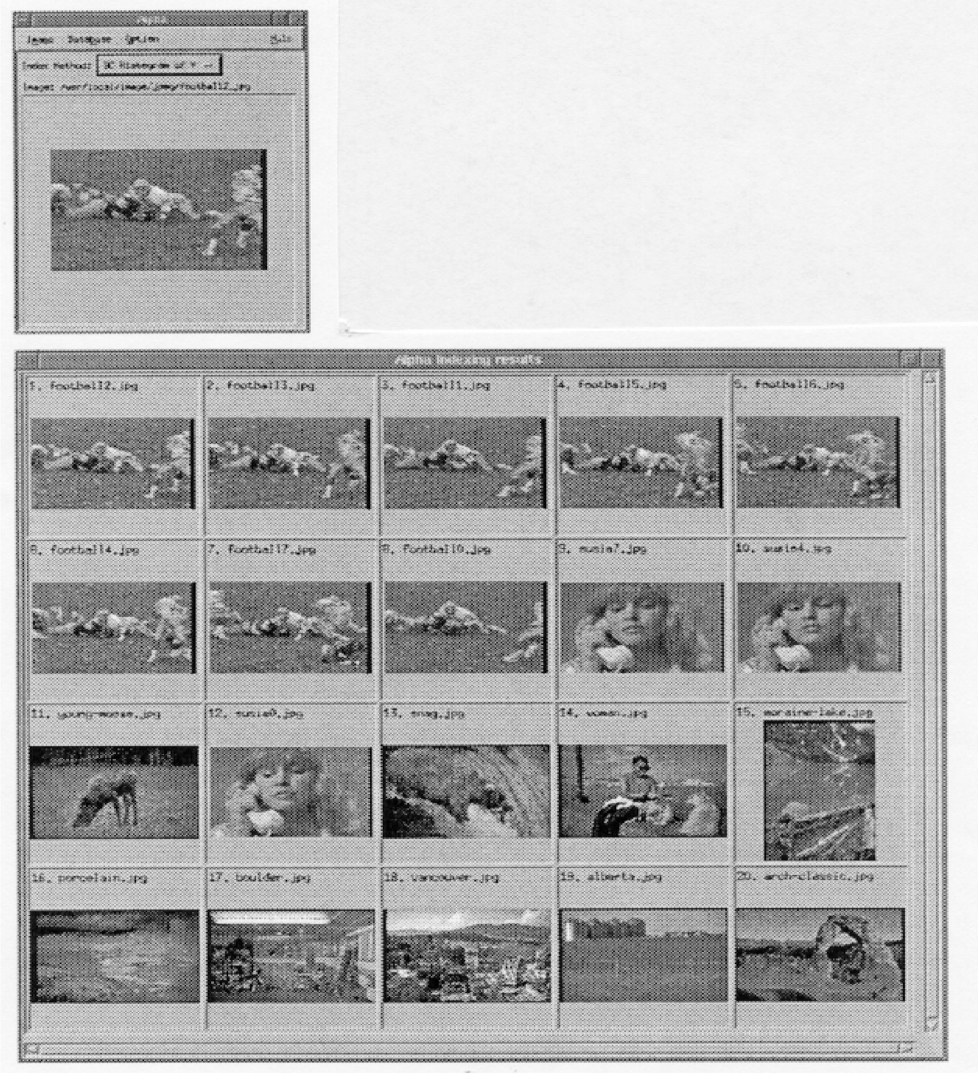
- Key frame extraction

- Indexing and retrieval of key frames

DC Histogram Technique Applied for Video Partitioning

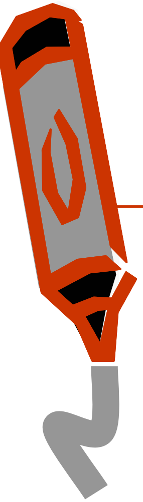


Example of Similarity-Based Retrieval of Key Frames Using DC Histograms



Index Method: DC Histogram of Y

Image: /usr/local/image/jpeg/football2.jpg



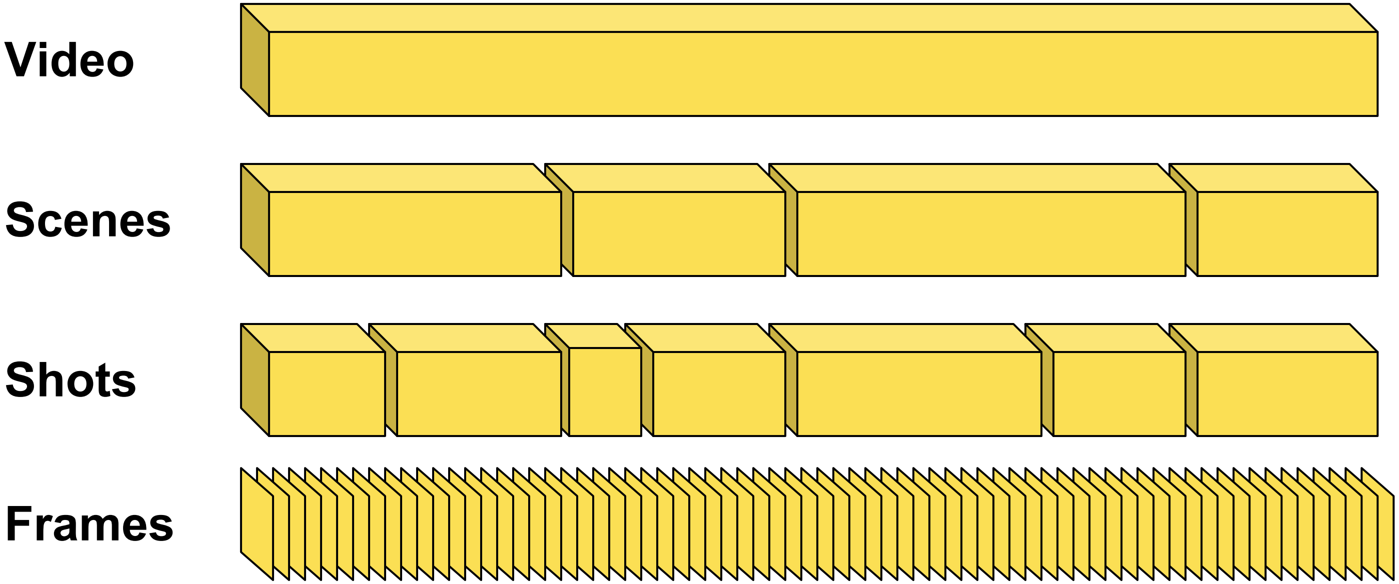
outline

- images
- video
- speech

Images and Video

- A digital image = a collection of pixels
  - Each pixel has a “color”
- Different types of pixels
  - Binary (1 bit): black/white
  - Grayscale (8 bits)
  - Color (3 colors, 8 bits each): red, green, blue
- A video is simply lots of images in rapid sequence
  - Each image is called a frame
  - Smooth motion requires about 24 frames/sec
- Compression is the key!

The Structure of Video



Video

Scenes

Shots

Frames

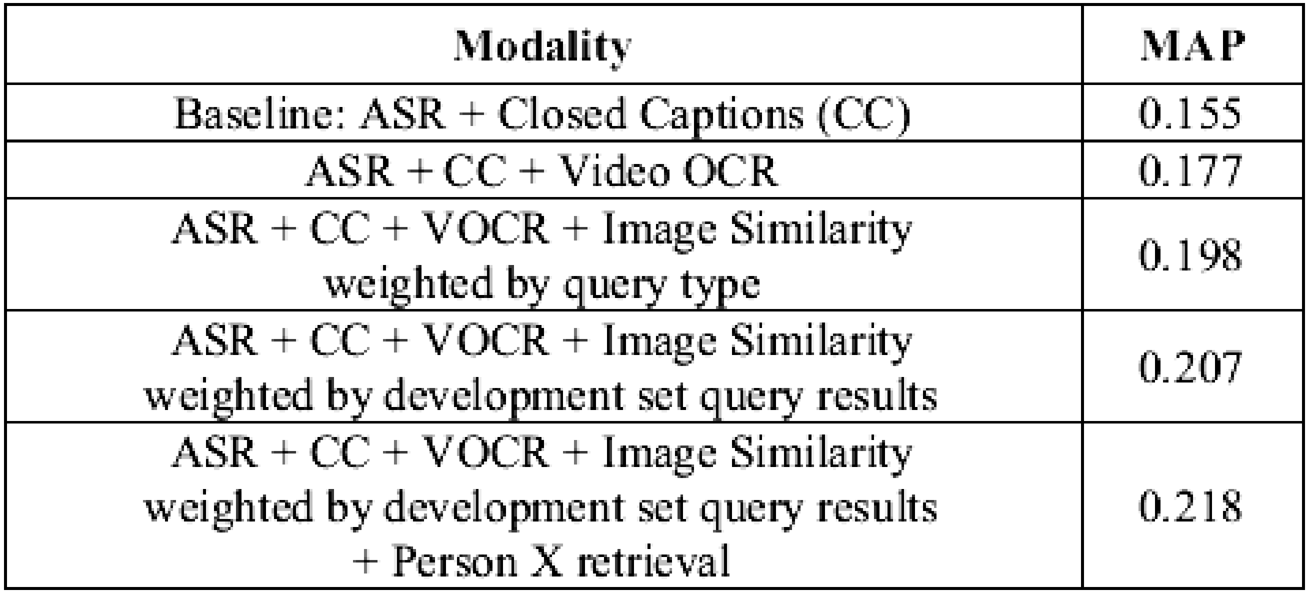
TREC For Video Retrieval?

- TREC Video Track (TRECVID)
  - Started in 2001
  - Goal is to investigate content-based retrieval from digital video
  - Focus on the shot as the unit of information retrieval
    (why?)

http://www-nlpir.nist.gov/projects/trecvid/

- Test Data Collection in 2004:
  - 74 hours of CNN Headline News, ABC World News Tonight, C-SPAN

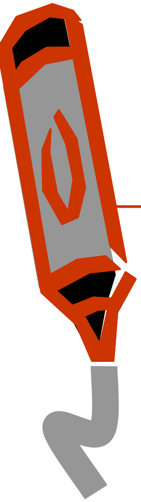
Searching Performance



Modality	MAP
Baseline: ASR + Closed Captions (CC)	0.155
ASR + CC + Video OCR	0.177
ASR + CC + VOCR + Image Similarity weighted by query type	0.198
ASR + CC + VOCR + Image Similarity weighted by development set query results	0.207
ASR + CC + VOCR + Image Similarity weighted by development set query results + Person X retrieval	0.218

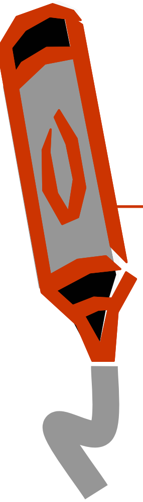
ASR = automatic speech recognition
CC = closed captions
VOCR = video optical character recognition

A. Hauptmann and M. Christel. (2004) Successful Approaches in the TREC Video Retrieval Evaluations. Proceedings of ACM Multimedia 2004.



Market Trends

- Broadband doubling over next 3-5 years
- Video enabled devices are emerging rapidly
- Emergence of mass internet audience
- Mainstream media moving to the Web
- International trends are similar
- Money Follows…



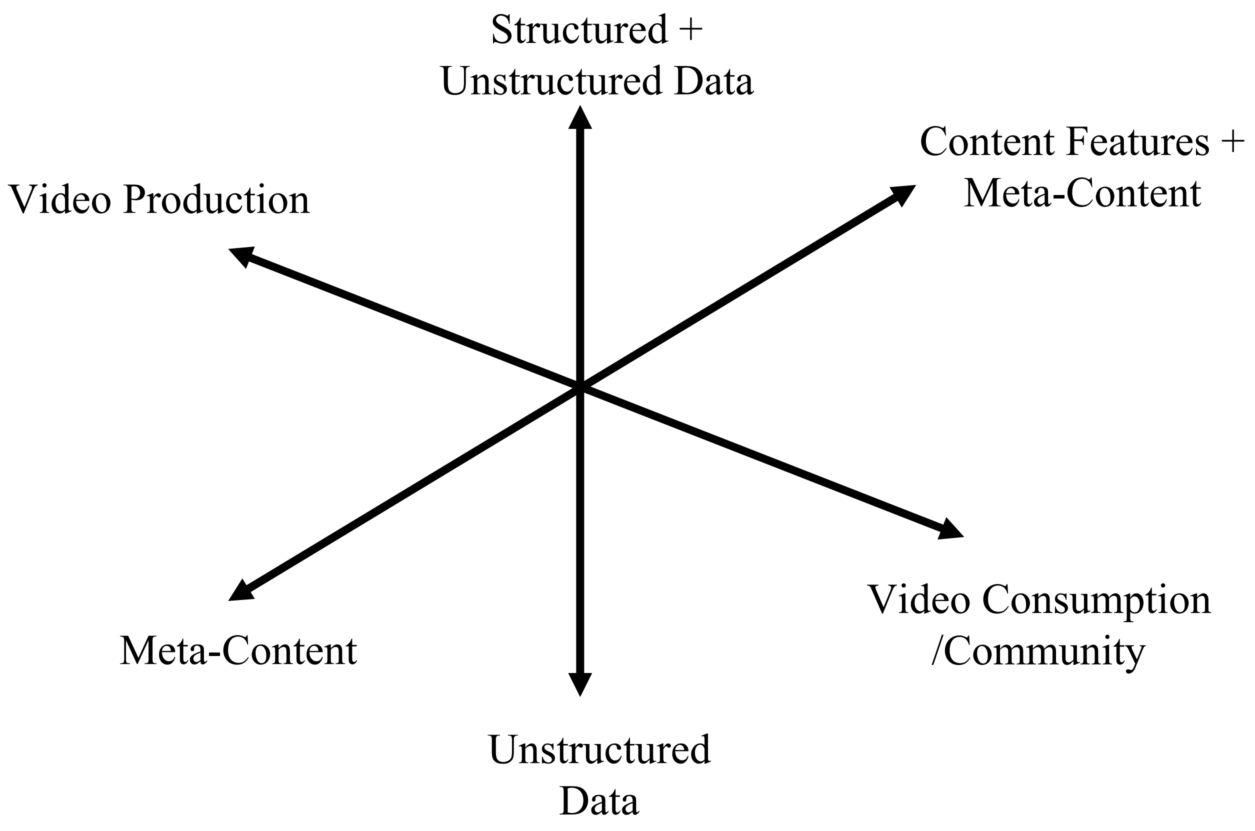
Market Trends(2005)

- How many of you are aware of video on the Web?
  - Large portion of online users
- How many have viewed a video on the Web in
  - The last 3 months?
    - 50%
  - The last 6 months?
  - Ever?
- Would you watch video on your devices (ipod/wireless)?
  - 1M downloads in 20 days (iPod)
- How many of you have produced video (personal or otherwise) recently?
  - Continuing to skyrocket with digital camera phones/devices
- How many of you have shared that with your friends/community? Would you have liked to?
  - Huge interest & adoption in viral communities

Market Trends

- Technology more media friendly
  - Storage costs plummeting (GB → TB)
  - CPU speed continuing to double (Moore's law)
  - Increased bandwidth
  - Device support for media
  - Adding media to sites drives traffic
  - Web continues to propel scalable infrastructure for media products/communities

Video Search



Structured + Unstructured Data

Video Production

Content Features + Meta-Content

Meta-Content

Unstructured Data

Video Consumption /Community

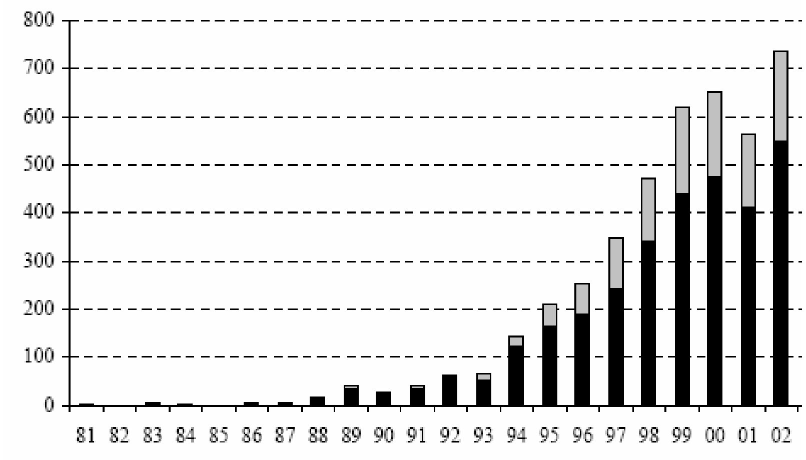
Video Search

- Media Information Retrieval
  - Been around since the late 1970s
    - Text Based/DB
      - Issues: Manual Annotation, Subjectivity of Human Perception
    - Content Based
      - Color, texture, shape, face detection/recognition, speech transcriptions, motion, segmentation boundaries/shots
      - High Dimensionality
      - Limited success to date.

Citations:
"Image Retrieval: Current Techniques, Promising Directions, and Open Issues" [Rui et al 99]

Video Search

Active Research Area



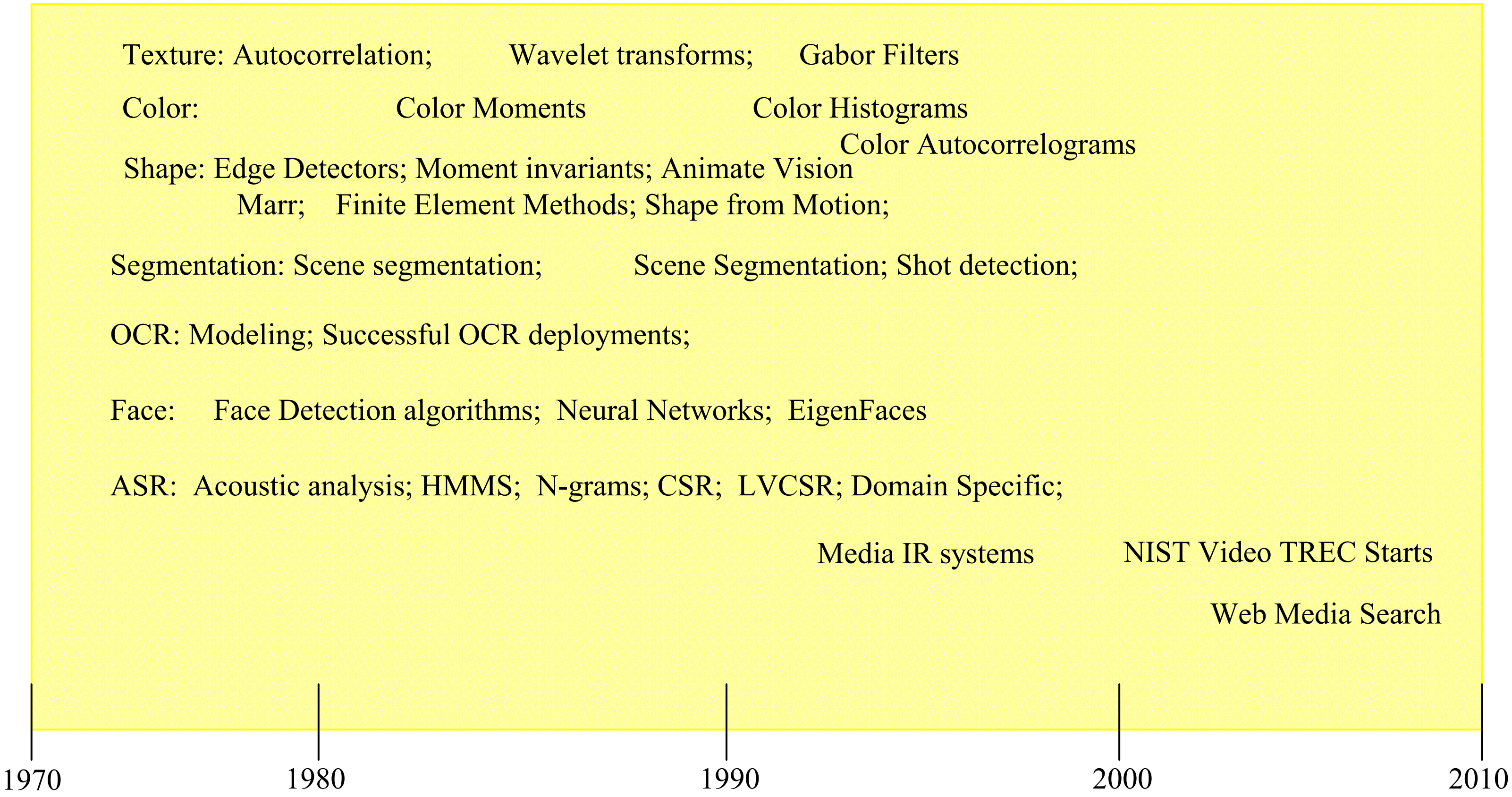
Graph from “A new perspective on Visual Information Retrieval”, Horst Eidenberger, 2004
Black: “Image Retrieval”; Grey:”Video Retrieval”; IEEE Digital Library

Video Search

Popular features/techniques:

- Color, Shape, Texture, Shape descriptors
- OCR, ASR
- A number of prototype or research products with small data sets
- More researched for visual queries

Video Search: Features



Texture: Autocorrelation; Wavelet transforms; Gabor Filters

Color: Color Moments Color Histograms

Color Autocorrelograms

Shape: Edge Detectors; Moment invariants; Animate Vision

Marr; Finite Element Methods; Shape from Motion;

Segmentation: Scene segmentation; Scene Segmentation; Shot detection;

OCR: Modeling; Successful OCR deployments;

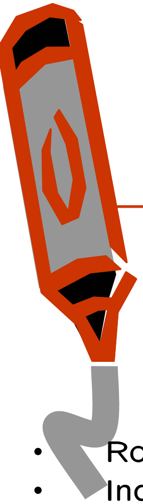
Face: Face Detection algorithms; Neural Networks; EigenFaces

ASR: Acoustic analysis; HMMS; N-grams; CSR; LVCSR; Domain Specific;

Media IR systems NIST Video TREC Starts

Web Media Search

Video Search: Features



Color

- Robust to background
- Independent of size, orientation
- Color Histogram [Swain & Ballard]
- “Sensitive to noise and sparse”- Cumulative Histograms [Stricker & Orgengo]
- Color Moments
- Color Sets: Map RGB Color space to Hue Saturation Value, & quantize [Smith, Chang]
- Color layout- local color features by dividing image into regions
- Color Autocorrelograms

Texture

- One of the earliest Image features [Harlick et al 70s]
- Co-occurrence matrix
- Orientation and distance on gray-scale pixels
- Contrast, inverse deference moment, and entropy [Gotlieb & Kreyszig]
- Human visual texture properties: coarseness, contrast, directionality, likeliness, regularity and roughness [Tamura et al]
- Wavelet Transforms [90s]
- [Smith & Chang] extracted mean and variance from wavelet subbands
- Gabor Filters
- And so on

Region Segmentation

- Partition image into regions
- Strong Segmentation: Object segmentation is difficult.
- Weak segmentation: Region segmentation based on some homegenity criteria

Scene Segmentation

- Shot detection, scene detection
- Look for changes in color, texture, brightness
- Context based scene segmentation applied to certain categories such as broadcast news

Video Search: Features

Shape

- Outer Boundary based vs. region based
- Fourier descriptors
- Moment invariants
- Finite Element Method (Stiffness matrix- how each point is connected to others; Eigen vectors of matrix)
- Turing function based (similar to Fourier descriptor) convex/concave polygons[Arkin et al]
- Wavelet transforms leverages multiresolution [Chuang & Kao]
- Chamfer matching for comparing 2 shapes (linear dimension rather than area)
- 3-D object representations using similar invariant features
- Well-known edge detection algorithms.

Face

- Face detection is highly reliable
  - Neural Networks [Rwoley]
  - Wavelet based histograms of facial features [Schneiderman]
- Face recognition for video is still a challenging problem.
  - EigenFaces: Extract eigenvectors and use as feature space

OCR

- OCR is fairly successful technology.
- Accurate, especially with good matching vocabularies.
- Script recognition still an open problem.

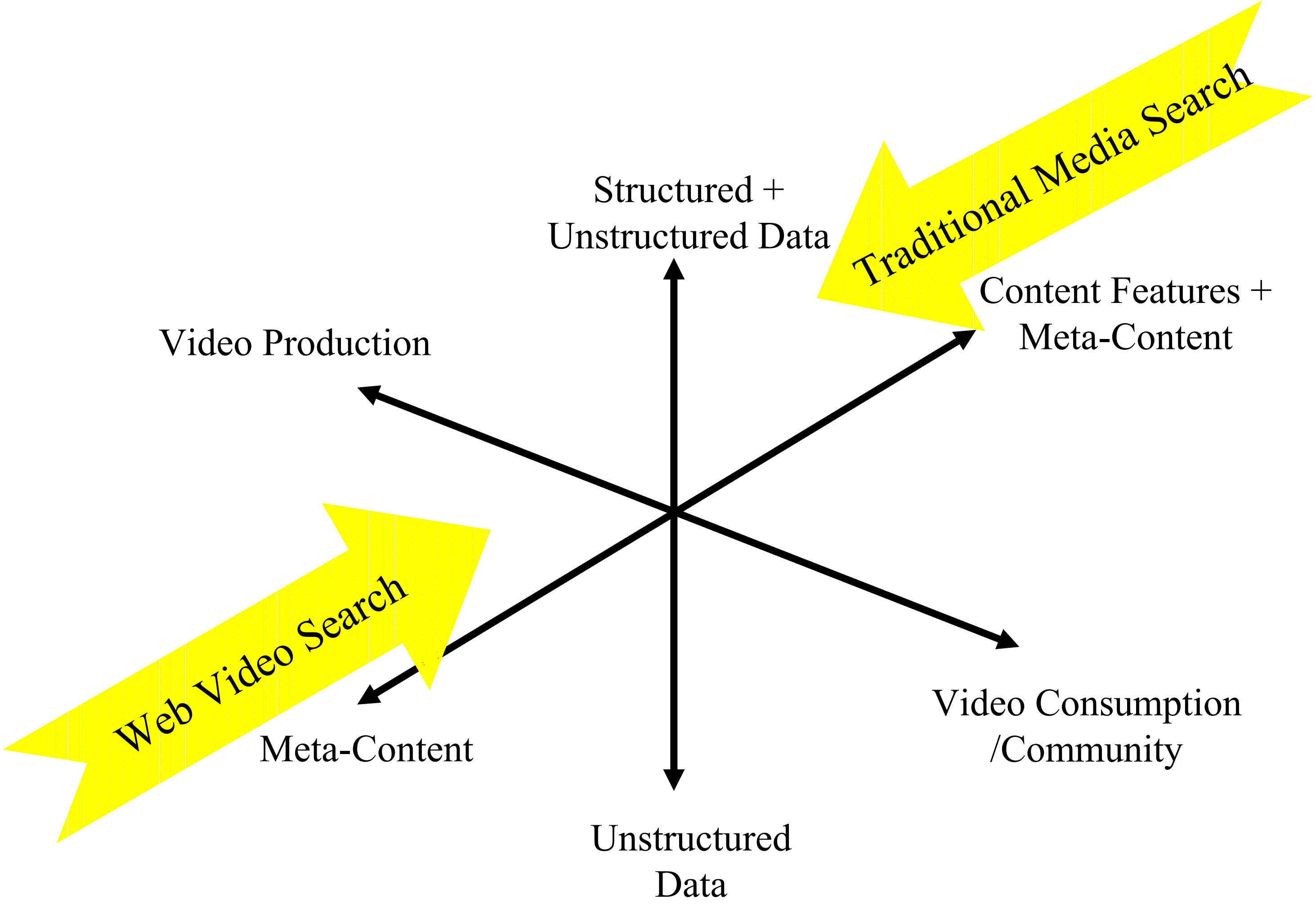
ASR

- Automatic speech recognition fairly accurate for medium to large vocabulary broadcast type data
- Large number of available speech vendors.
- Still open for free conversational speech in noisy conditions.

Video Search: Video TREC

- Overview:
  - Shot detection, story segmentation, semantic feature extraction, information retrieval
  - Corpora of documentaries, advertising films
  - Broadcast news added in 2003
  - Interactive and non-interactive tests
  - CBIR features
  - Speech transcribed (LIMSI)
  - OCR

Video Search



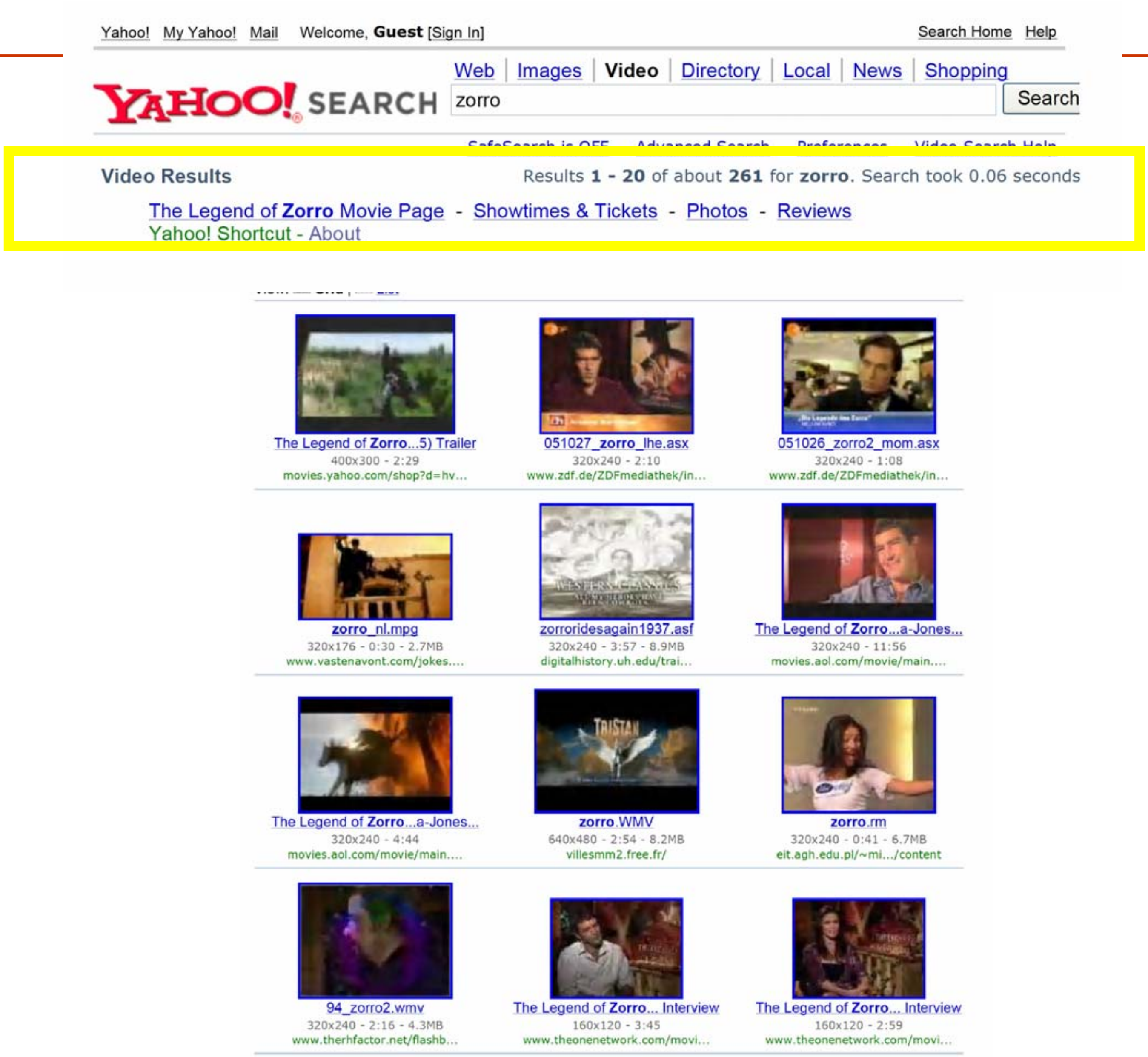
Opportunities

Example 1:

- User query “zorro”

- User 1 wants to see Zorro videos
- User 2 wants to see Legend of Zorro movie clips
- User 3 just wants to see home videos about Zorro

- Can content based analysis help over structured meta-data query inference?



Yahoo! My Yahoo! Mail Welcome, **Guest** [Sign In] Search Home Help

Web | Images | **Video** | Directory | Local | News | Shopping

YAHOO! SEARCH zorro Search

SafeSearch is OFF Advanced Search Preferences Video Search Help

Video Results Results **1 - 20** of about **261** for **zorro**. Search took 0.06 seconds

The Legend of **Zorro** Movie Page - Showtimes & Tickets - Photos - Reviews
Yahoo! Shortcut - About

The Legend of **Zorro**...5) Trailer
400x300 - 2:29
movies.yahoo.com/shop?d=hv...

051027_**zorro**_lhe.asx
320x240 - 2:10
www.zdf.de/ZDFmediathek/in...

051026_zorro2_mom.asx
320x240 - 1:08
www.zdf.de/ZDFmediathek/in...

zorro_nl.mpg
320x176 - 0:30 - 2.7MB
www.vastenavont.com/jokes....

zorroridesagain1937.asf
320x240 - 3:57 - 8.9MB
digitalhistory.uh.edu/trai...

The Legend of **Zorro**...a-Jones...
320x240 - 11:56
movies.aol.com/movie/main....

The Legend of **Zorro**...a-Jones...
320x240 - 4:44
movies.aol.com/movie/main....

zorro.WMV
640x480 - 2:54 - 8.2MB
villesmm2.free.fr/

zorro.rm
320x240 - 0:41 - 6.7MB
eit.agh.edu.pl/~mi.../content

94_zorro2.wmv
320x240 - 2:16 - 4.3MB
www.therhfactor.net/flashb...

The Legend of **Zorro**... Interview
160x120 - 3:45
www.theonenetwork.com/movi...

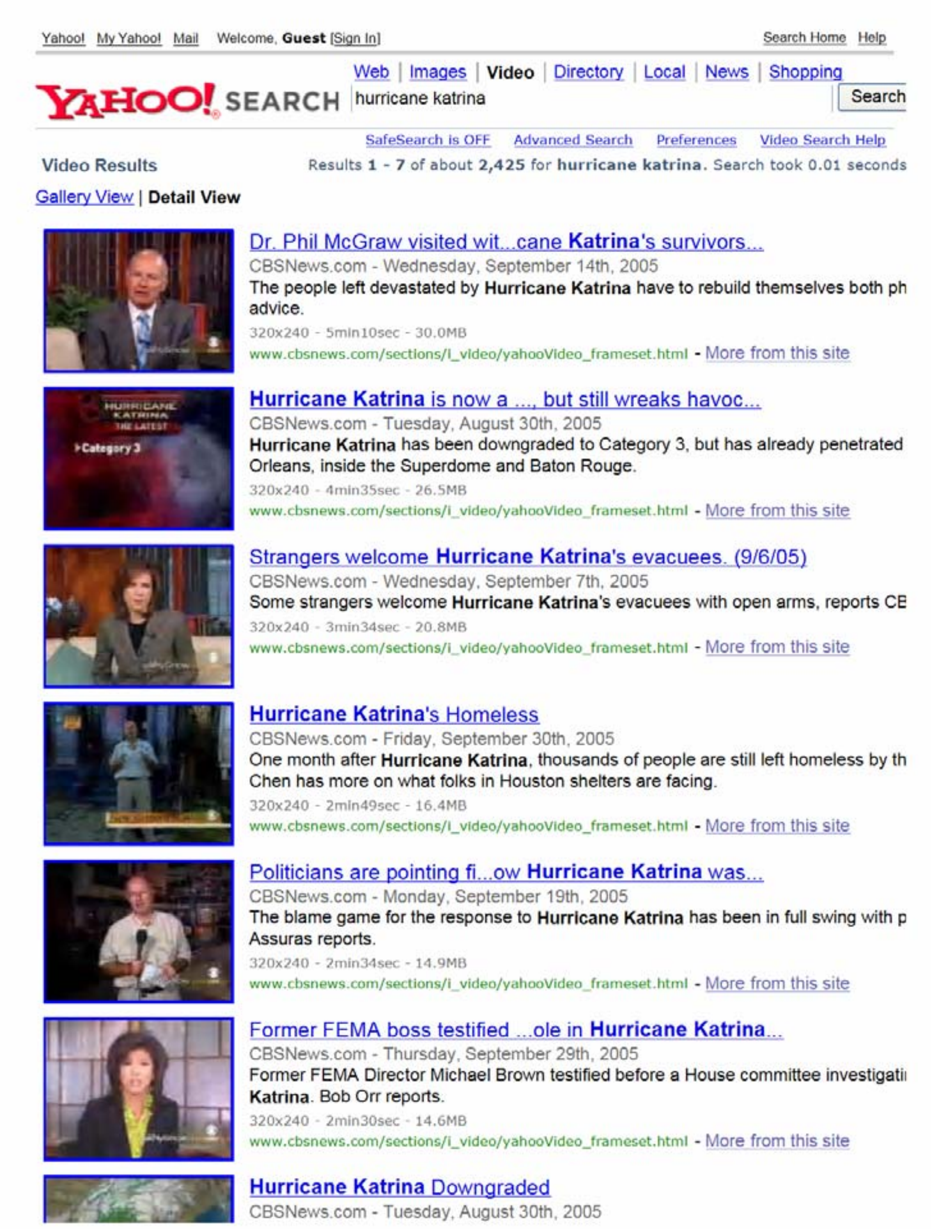
The Legend of **Zorro**... Interview
160x120 - 2:59
www.theonenetwork.com/movi...

Opportunities

Example 2:

- For main-stream head content such as news videos.
- Meta-data are fairly descriptive
- Usually queried based on non-visual attributes.

- Task: “Pull up recent Hurricane Katrina videos”



Welcome, **Guest** [Sign In]

hurricane katrina

Video Results

Results **1 - 7** of about **2,425** for **hurricane katrina**. Search took 0.01 seconds

Dr. Phil McGraw visited wit...cane **Katrina**'s survivors...
CBSNews.com - Wednesday, September 14th, 2005
The people left devastated by **Hurricane Katrina** have to rebuild themselves both ph
advice.
320x240 - 5min10sec - 30.0MB
www.cbsnews.com/sections/i_video/yahooVideo_frameset.html - More from this site

Hurricane Katrina is now a ..., but still wreaks havoc...
CBSNews.com - Tuesday, August 30th, 2005
Hurricane Katrina has been downgraded to Category 3, but has already penetrated
Orleans, inside the Superdome and Baton Rouge.
320x240 - 4min35sec - 26.5MB
www.cbsnews.com/sections/i_video/yahooVideo_frameset.html - More from this site

Strangers welcome **Hurricane Katrina**'s evacuees. (9/6/05)
CBSNews.com - Wednesday, September 7th, 2005
Some strangers welcome **Hurricane Katrina**'s evacuees with open arms, reports CE
320x240 - 3min34sec - 20.8MB
www.cbsnews.com/sections/i_video/yahooVideo_frameset.html - More from this site

Hurricane Katrina's Homeless
CBSNews.com - Friday, September 30th, 2005
One month after **Hurricane Katrina**, thousands of people are still left homeless by th
Chen has more on what folks in Houston shelters are facing.
320x240 - 2min49sec - 16.4MB
www.cbsnews.com/sections/i_video/yahooVideo_frameset.html - More from this site

Politicians are pointing fi...ow **Hurricane Katrina** was...
CBSNews.com - Monday, September 19th, 2005
The blame game for the response to **Hurricane Katrina** has been in full swing with p
Assuras reports.
320x240 - 2min34sec - 14.9MB
www.cbsnews.com/sections/i_video/yahooVideo_frameset.html - More from this site

Former FEMA boss testified ...ole in **Hurricane Katrina**...
CBSNews.com - Thursday, September 29th, 2005
Former FEMA Director Michael Brown testified before a House committee investigati
Katrina. Bob Orr reports.
320x240 - 2min30sec - 14.6MB
www.cbsnews.com/sections/i_video/yahooVideo_frameset.html - More from this site

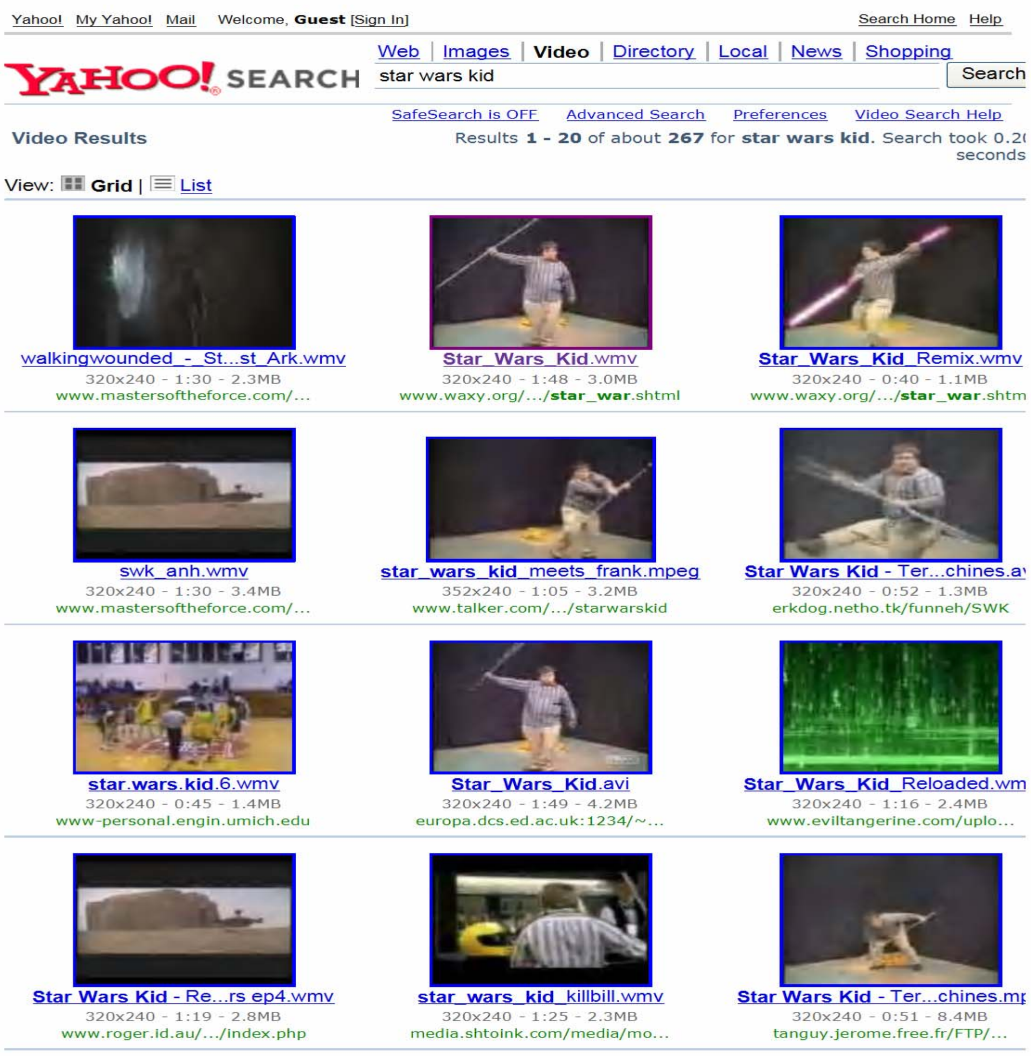
Hurricane Katrina Downgraded
CBSNews.com - Tuesday, August 30th, 2005

Opportunities

Example 3:

- Creative Home Video
- Community video rendering!

- The now “famous” Star Wars Kid
- Example of “social buzz” combined with innovative tail content video production.



Yahoo! My Yahoo! Mail Welcome, **Guest** [Sign In] Search Home Help

Web | Images | **Video** | Directory | Local | News | Shopping

YAHOO! SEARCH star wars kid Search

SafeSearch is OFF Advanced Search Preferences Video Search Help

Video Results Results **1 - 20** of about **267** for **star wars kid**. Search took 0.2(seconds

View: **Grid** | List

walkingwounded_-_St...st_Ark.wmv
320x240 - 1:30 - 2.3MB
www.mastersoftheforce.com/...

Star_Wars_Kid.wmv
320x240 - 1:48 - 3.0MB
www.waxy.org/.../**star_war**.shtml

Star_Wars_Kid_Remix.wmv
320x240 - 0:40 - 1.1MB
www.waxy.org/.../**star_war**.shtm

swk_anh.wmv
320x240 - 1:30 - 3.4MB
www.mastersoftheforce.com/...

star_wars_kid_meets_frank.mpeg
352x240 - 1:05 - 3.2MB
www.talker.com/.../starwarskid

Star Wars Kid - Ter...chines.a\
320x240 - 0:52 - 1.3MB
erkdog.netho.tk/funneh/SWK

star.wars.kid.6.wmv
320x240 - 0:45 - 1.4MB
www-personal.engin.umich.edu

Star_Wars_Kid.avi
320x240 - 1:49 - 4.2MB
europa.dcs.ed.ac.uk:1234/~...

Star_Wars_Kid_Reloaded.wm
320x240 - 1:16 - 2.4MB
www.eviltangerine.com/uplo...

Star Wars Kid - Re...rs ep4.wmv
320x240 - 1:19 - 2.8MB
www.roger.id.au/.../index.php

star_wars_kid_killbill.wmv
320x240 - 1:25 - 2.3MB
media.shtoink.com/media/mo...

Star Wars Kid - Ter...chines.mp
320x240 - 0:51 - 8.4MB
tanguy.jerome.free.fr/FTP/...

Play Video 1

Play Video 2

Opportunities

A hard example:

- Lets take an example:

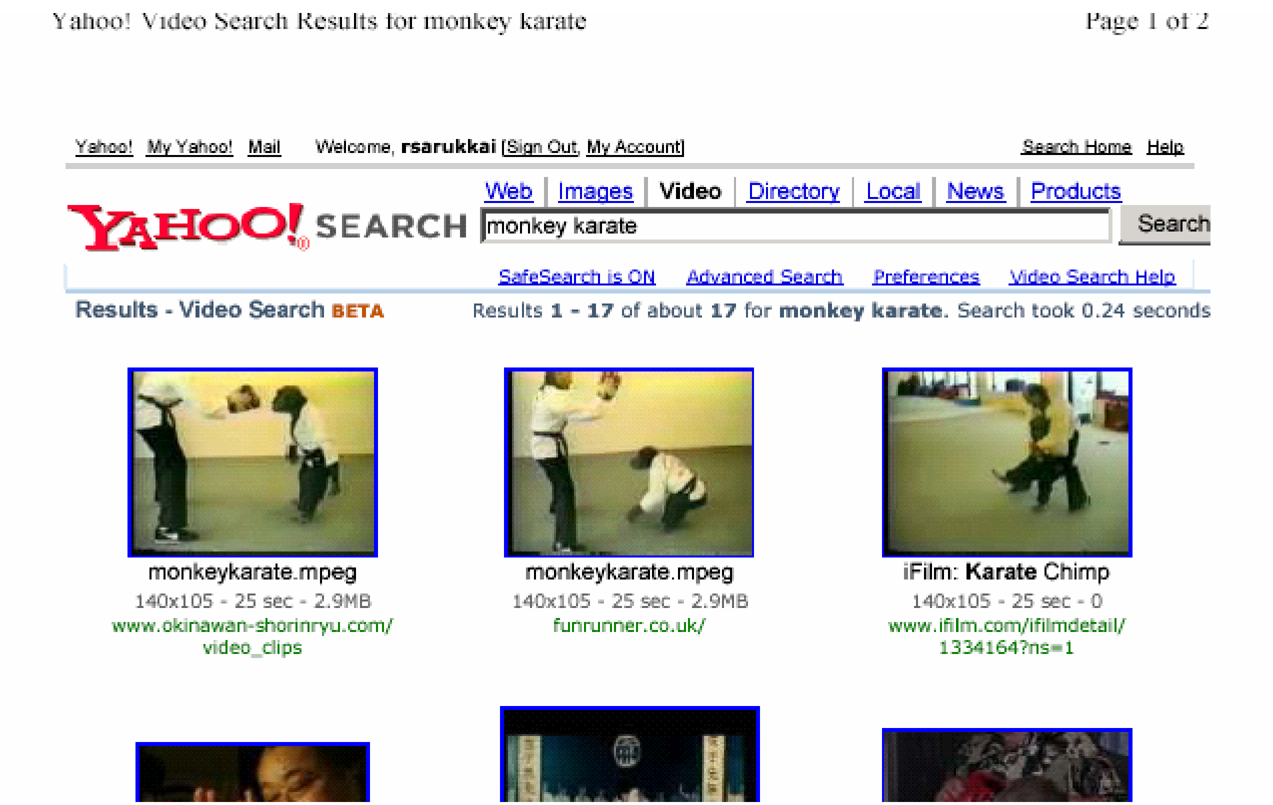
  “Supposing you want to find videos that depict a monkey/chimp doing karate”!

- CBIR Approach:

  Train models for Chimps/Monkeys

  Motion Analysis for Karate movement models

  Many open issues/problems!



Yahoo! Video Search Results for monkey karate   Page 1 of 2

Yahoo! My Yahoo! Mail   Welcome, **rsarukkai** [Sign Out, My Account]   Search Home Help

YAHOO! SEARCH
Web | Images | Video | Directory | Local | News | Products
monkey karate
Search
SafeSearch is ON   Advanced Search   Preferences   Video Search Help

Results - Video Search BETA   Results **1 - 17** of about **17** for **monkey karate**. Search took 0.24 seconds

monkeykarate.mpeg
140x105 - 25 sec - 2.9MB
www.okinawan-shorinryu.com/video_clips

monkeykarate.mpeg
140x105 - 25 sec - 2.9MB
funrunner.co.uk/

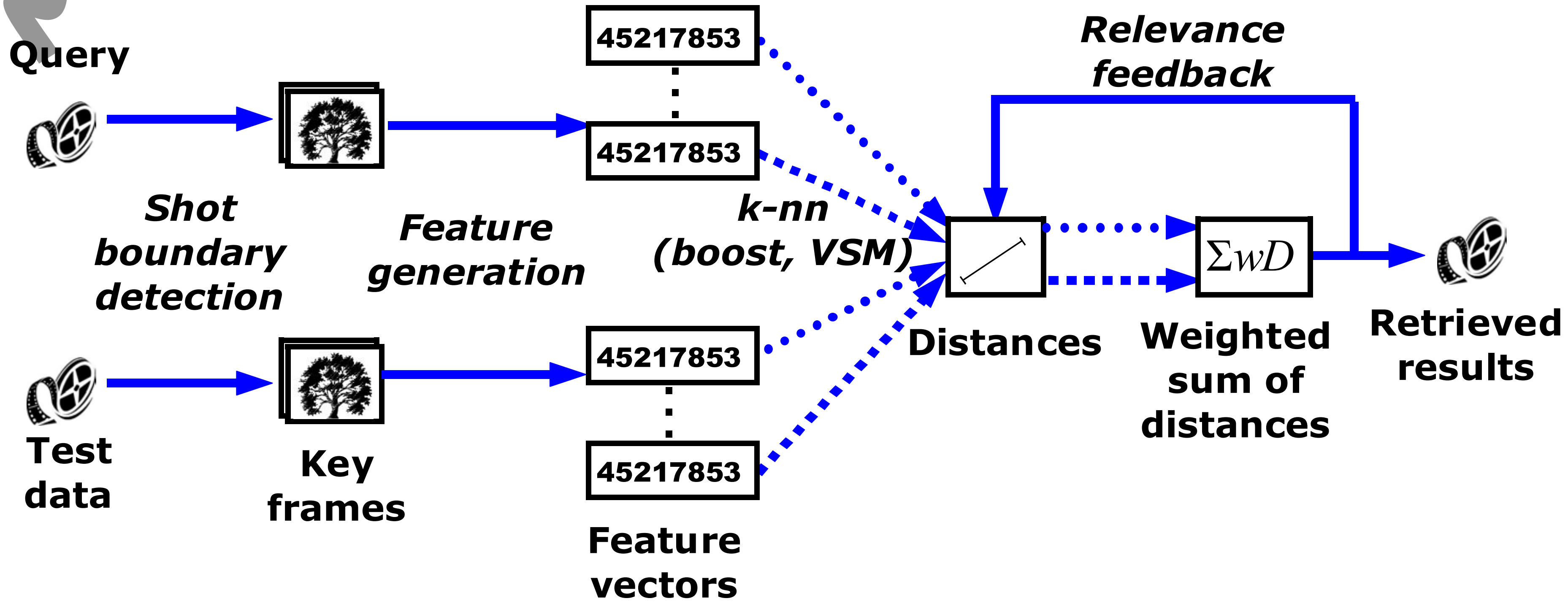
iFilm: **Karate** Chimp
140x105 - 25 sec - 0
www.ifilm.com/ifilmdetail/1334164?ns=1

Play Video

Video Data Management

1. Video Parsing
   - Manipulation of whole video for breakdown into key frames.
     - Scene: single dramatic event taken by a small number of related cameras.
     - Shot: A sequence taken by a single camera
     - Frame: A still image
2. Video Indexing
   - Retrieving information about the frame for indexing in a database.
3. Video Retrieval and browsing
   - Users access the db through queries or through interactions.

System overview



Query

Test data

Shot boundary detection

Key frames

Feature generation

45217853

45217853

45217853

45217853

Feature vectors

k-nn (boost, VSM)

Distances

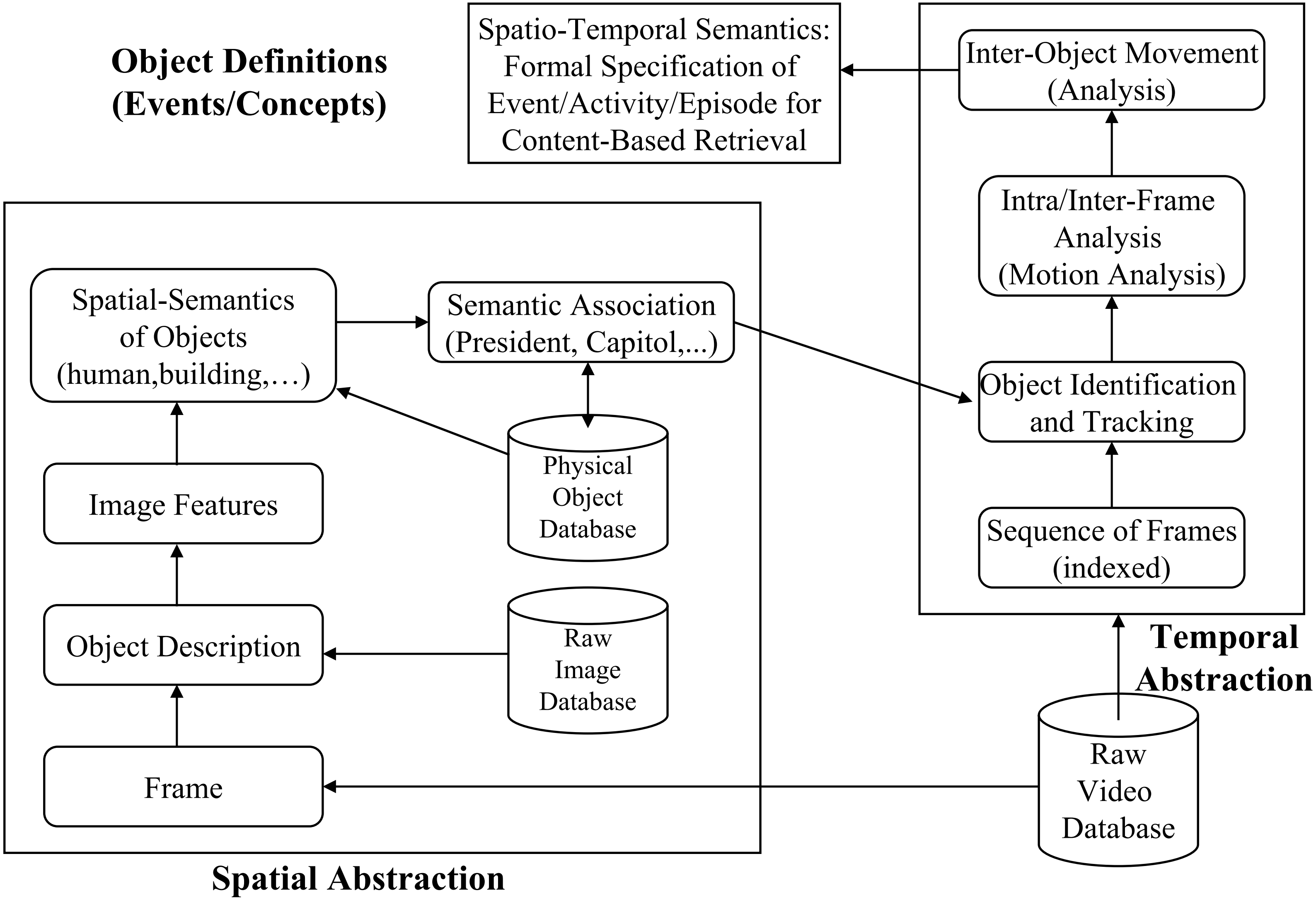
Relevance feedback

ΣwD

Weighted sum of distances

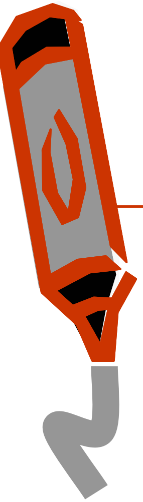
Retrieved results

An Architecture for Video Database System


Object Definitions
(Events/Concepts)
Spatio-Temporal Semantics: Formal Specification of Event/Activity/Episode for Content-Based Retrieval
Inter-Object Movement (Analysis)
Intra/Inter-Frame Analysis (Motion Analysis)
Object Identification and Tracking
Sequence of Frames (indexed)
Temporal Abstraction
Spatial-Semantics of Objects (human,building,…)
Semantic Association (President, Capitol,...)
Image Features
Physical Object Database
Object Description
Raw Image Database
Frame
Raw Video Database
Spatial Abstraction

Video Data Management

- Metadata-based method
- Text-based method
- Audio-based method
- Content-based method
- Integrated approach



Metadata-based Method

- Video is indexed and retrieved based on structured metadata information by using a traditional DBMS

- Metadata examples are the title, author, producer, director, date, types of video.

Text-based Method

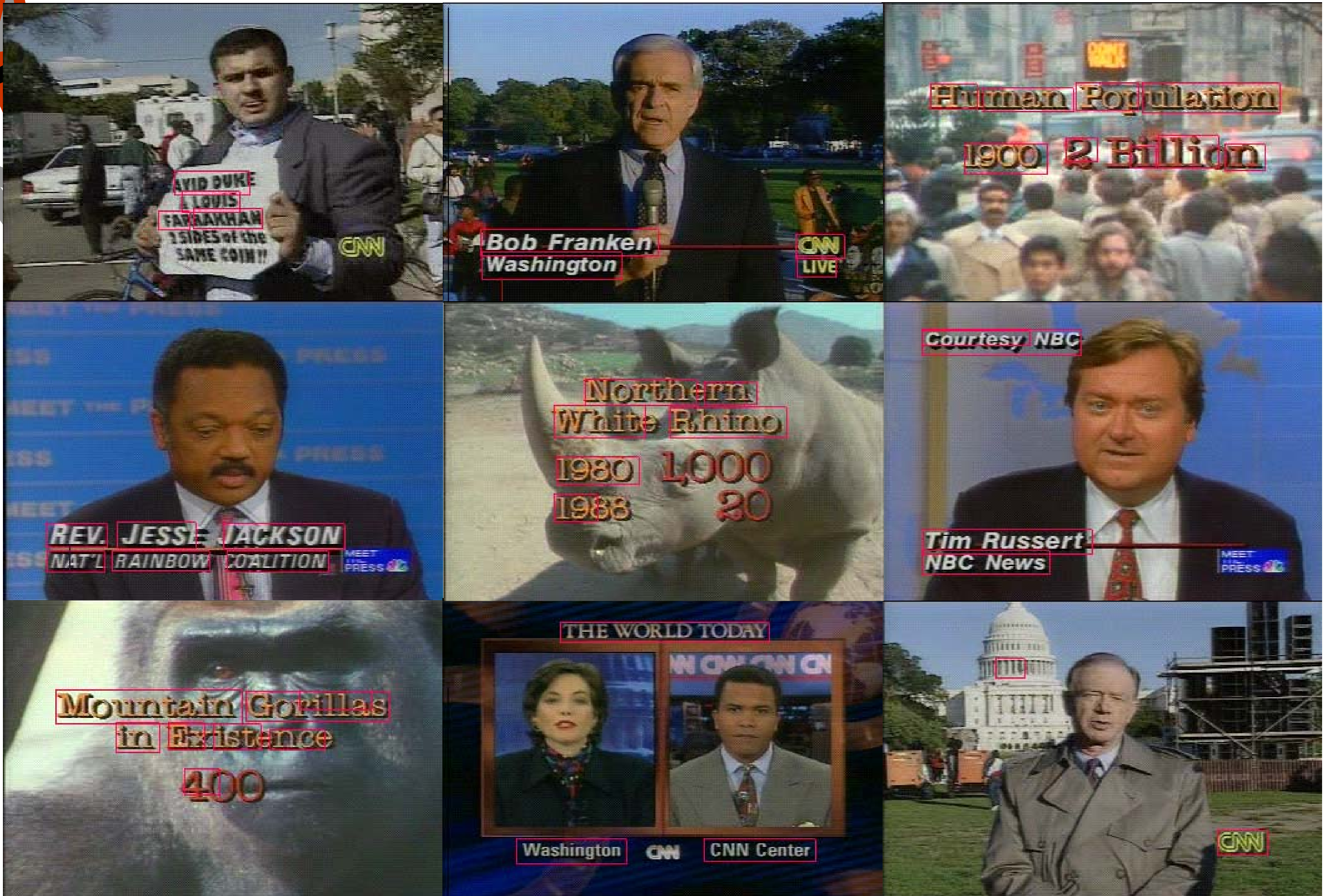
- Video is indexed and retrieved based on associated subtitles (text) using traditional IR techniques for text documents.

- Transcripts and subtitles are already exist in many types of video such as news and movies, eliminating the need for manual annotation.

Text-based Method

- Basic method is to use human annotation
- Can be done automatically where subtitles / transcriptions exist
  - BBC: 100% output subtitled by 2008
- Speech recognition for archive material

Text Detection

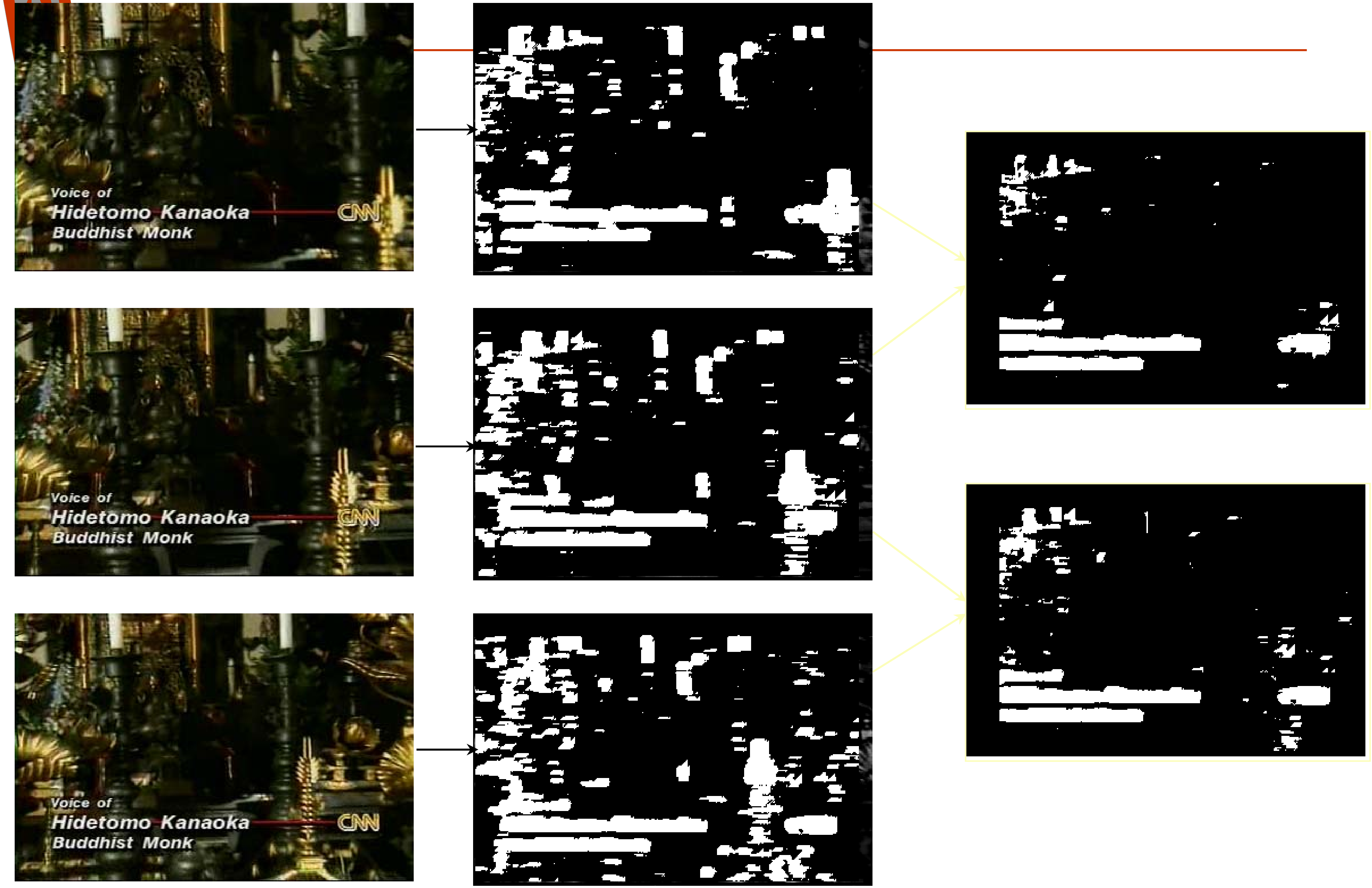


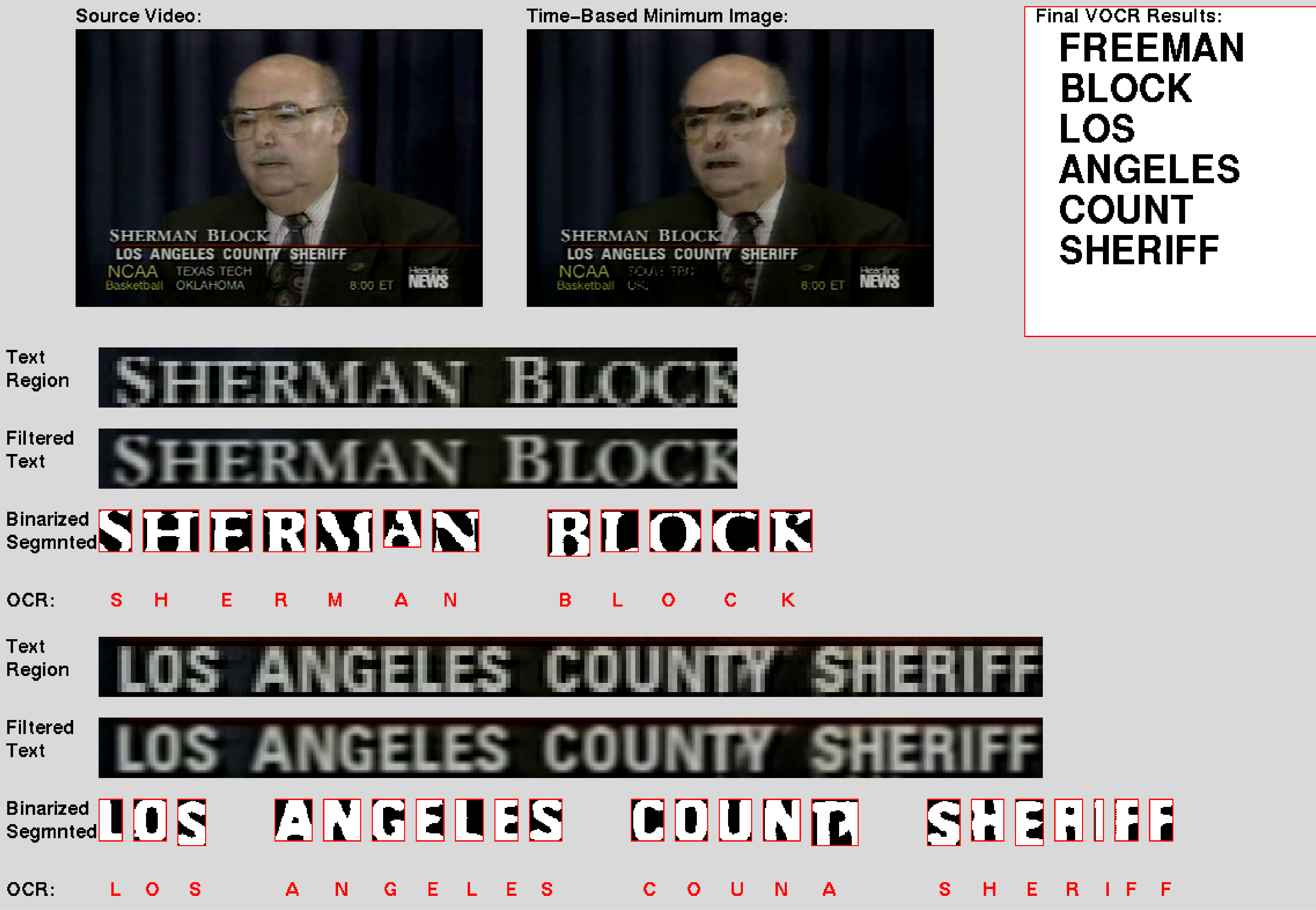
Video Frames

(1/2 s intervals)

Filtered Frames

AND-ed Frames





Source Video:

Time-Based Minimum Image:

Final VOCR Results:

FREEMAN
BLOCK
LOS
ANGELES
COUNT
SHERIFF

Text Region

Filtered Text

Binarized Segmnted

OCR: S H E R M A N B L O C K

Text Region

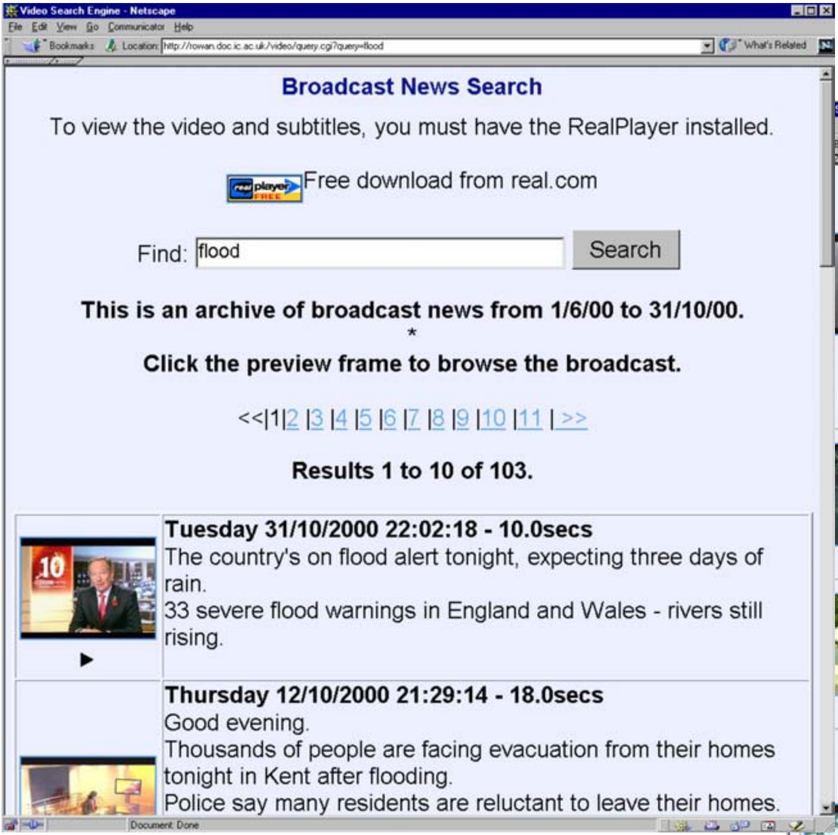
Filtered Text

Binarized Segmnted

OCR: L O S A N G E L E S C O U N A S H E R I F F

Text-based Method

- Key word search based on subtitles
- Content based


Video Search Engine - Netscape
File Edit View Go Communicator Help
Bookmarks Location: http://rowan.doc.ic.ac.uk/video/query.cgi?query=flood
What's Related
Document: Done

Broadcast News Search

To view the video and subtitles, you must have the RealPlayer installed.

Free download from real.com

Find: flood Search

This is an archive of broadcast news from 1/6/00 to 31/10/00.

*

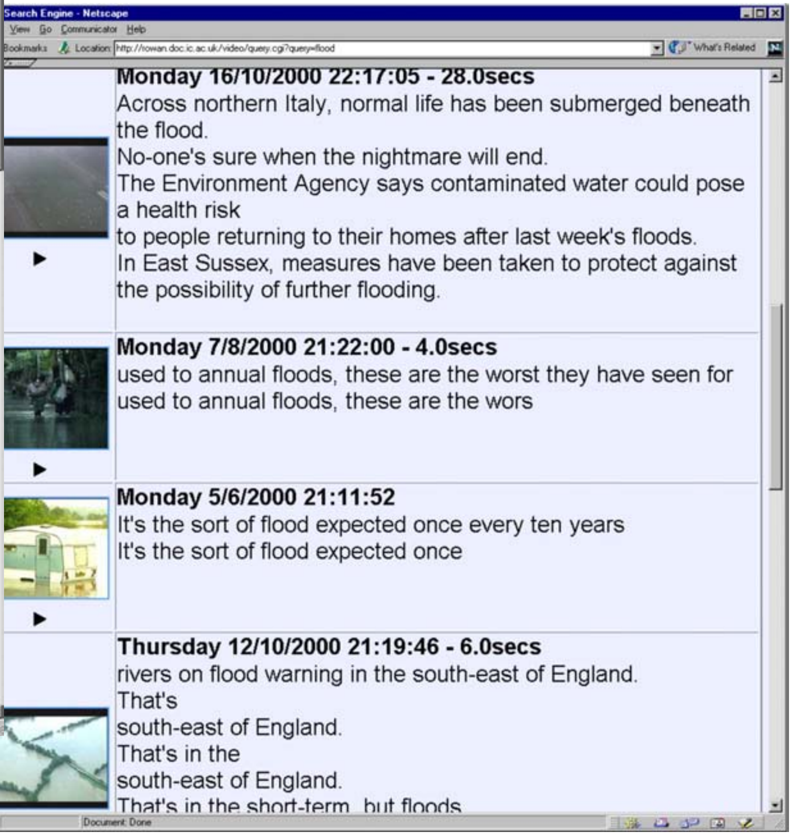
Click the preview frame to browse the broadcast.

<<|1|2 |3 |4 |5 |6 |7 |8 |9 |10 |11 | >>

Results 1 to 10 of 103.

Tuesday 31/10/2000 22:02:18 - 10.0secs
The country's on flood alert tonight, expecting three days of rain.
33 severe flood warnings in England and Wales - rivers still rising.

Thursday 12/10/2000 21:29:14 - 18.0secs
Good evening.
Thousands of people are facing evacuation from their homes tonight in Kent after flooding.
Police say many residents are reluctant to leave their homes.


Search Engine - Netscape
View Go Communicator Help
Bookmarks Location: http://rowan.doc.ic.ac.uk/video/query.cgi?query=flood
What's Related
Document: Done

Monday 16/10/2000 22:17:05 - 28.0secs
Across northern Italy, normal life has been submerged beneath the flood.
No-one's sure when the nightmare will end.
The Environment Agency says contaminated water could pose a health risk
to people returning to their homes after last week's floods.
In East Sussex, measures have been taken to protect against the possibility of further flooding.

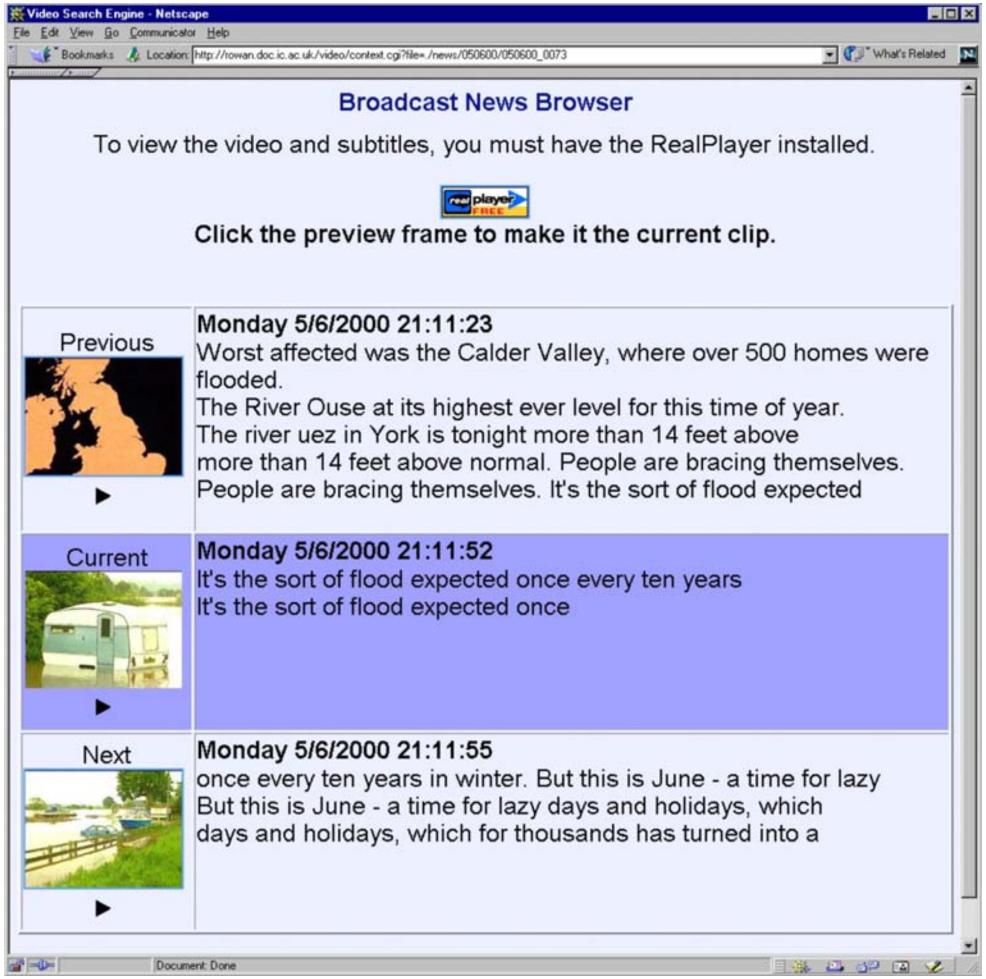
Monday 7/8/2000 21:22:00 - 4.0secs
used to annual floods, these are the worst they have seen for
used to annual floods, these are the wors

Monday 5/6/2000 21:11:52
It's the sort of flood expected once every ten years
It's the sort of flood expected once

Thursday 12/10/2000 21:19:46 - 6.0secs
rivers on flood warning in the south-east of England.
That's
south-east of England.
That's in the
south-east of England.
That's in the short-term, but floods

- Live demo:
  http://km.doc.ic.ac.uk/vse/

Text-based Method



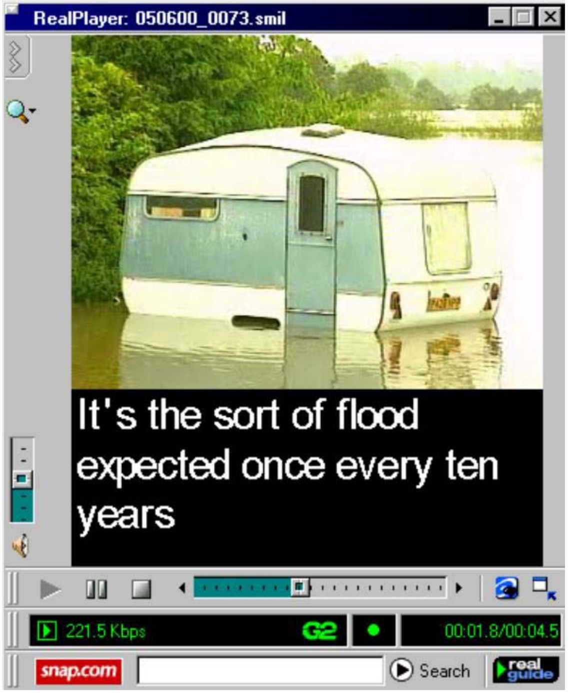
Video Search Engine - Netscape
File Edit View Go Communicator Help
Bookmarks
Location: http://rowan.doc.ic.ac.uk/video/context.cgi?file=./news/050600/050600_0073
What's Related
real player FREE
Document: Done

Broadcast News Browser

To view the video and subtitles, you must have the RealPlayer installed.

Click the preview frame to make it the current clip.

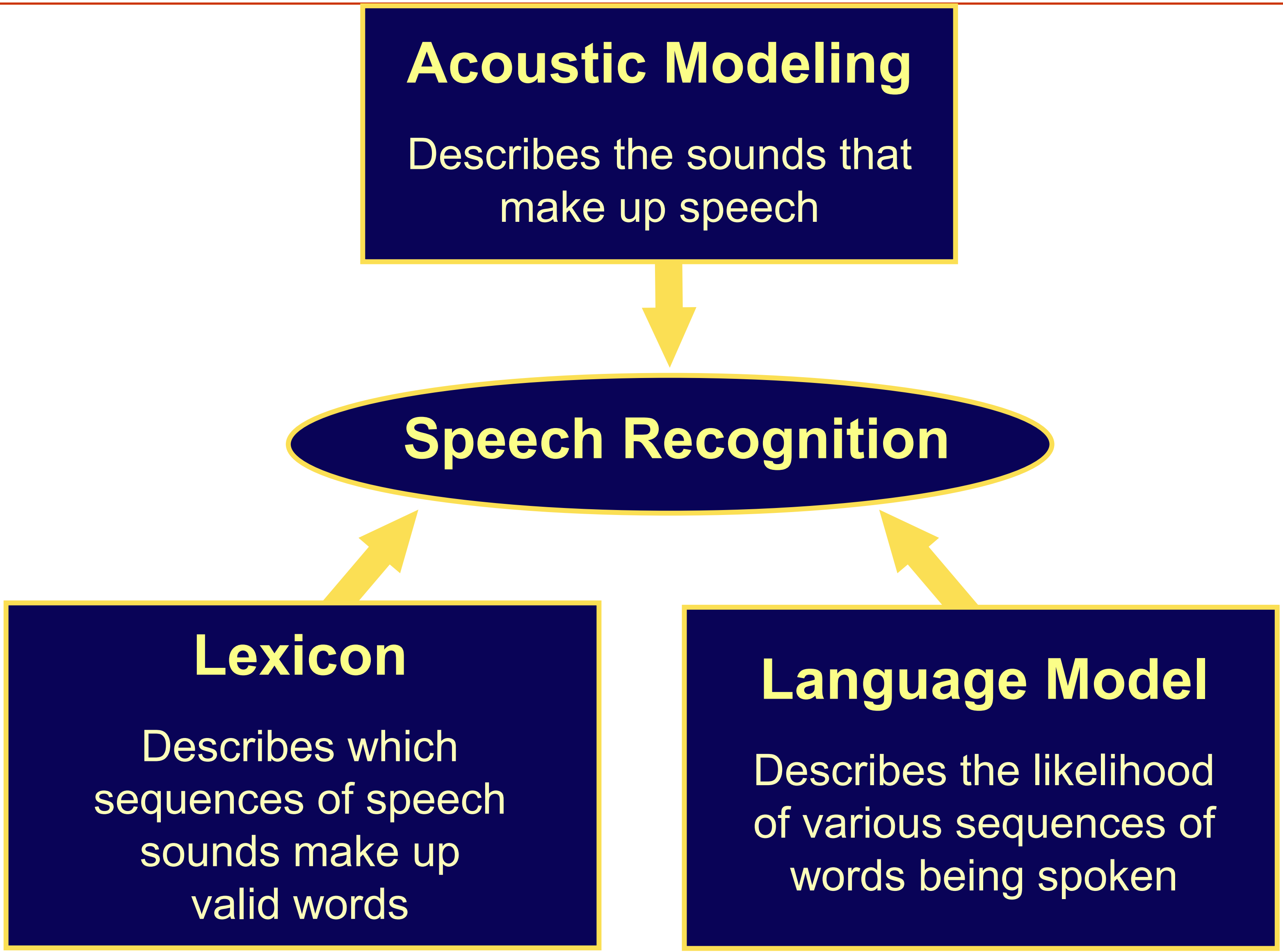
Previous ▶	**Monday 5/6/2000 21:11:23** Worst affected was the Calder Valley, where over 500 homes were flooded. The River Ouse at its highest ever level for this time of year. The river uez in York is tonight more than 14 feet above more than 14 feet above normal. People are bracing themselves. People are bracing themselves. It's the sort of flood expected
Current ▶	**Monday 5/6/2000 21:11:52** It's the sort of flood expected once every ten years It's the sort of flood expected once
Next ▶	**Monday 5/6/2000 21:11:55** once every ten years in winter. But this is June - a time for lazy But this is June - a time for lazy days and holidays, which days and holidays, which for thousands has turned into a



It's the sort of flood expected once every ten years

outline

- images
- video
- speech



Acoustic Modeling

Describes the sounds that make up speech

Speech Recognition

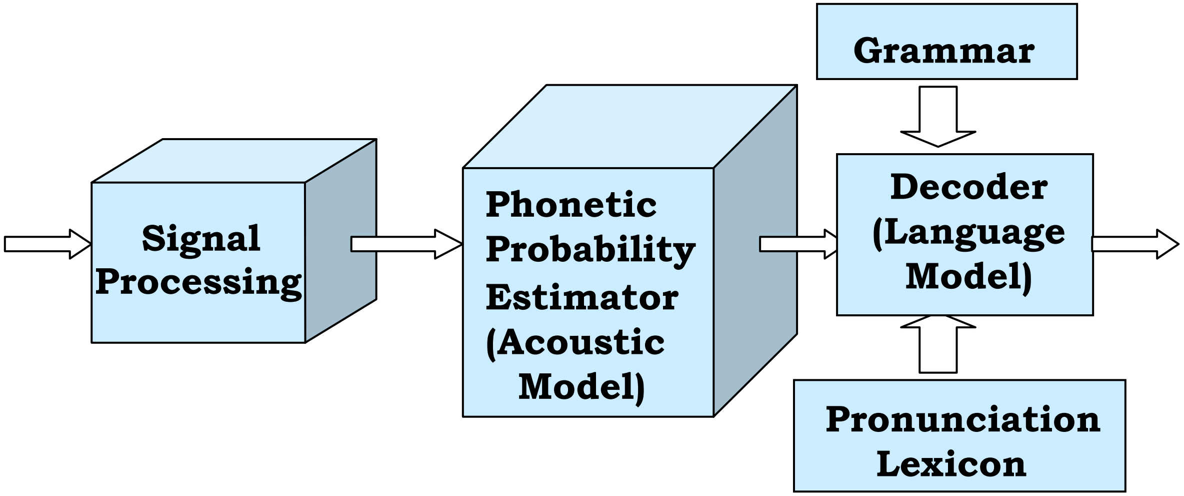
Lexicon

Describes which sequences of speech sounds make up valid words

Language Model

Describes the likelihood of various sequences of words being spoken

Speech Recognition in Brief



Hints For Better Recognition

- Goal: improve the estimation p(word|acoustic_sig)
- Main idea:

  p(word|acoustic_sign) → p(word|acoustic_signal, X)

What could be X?

- Topical information
- News of the day
- Image information ?

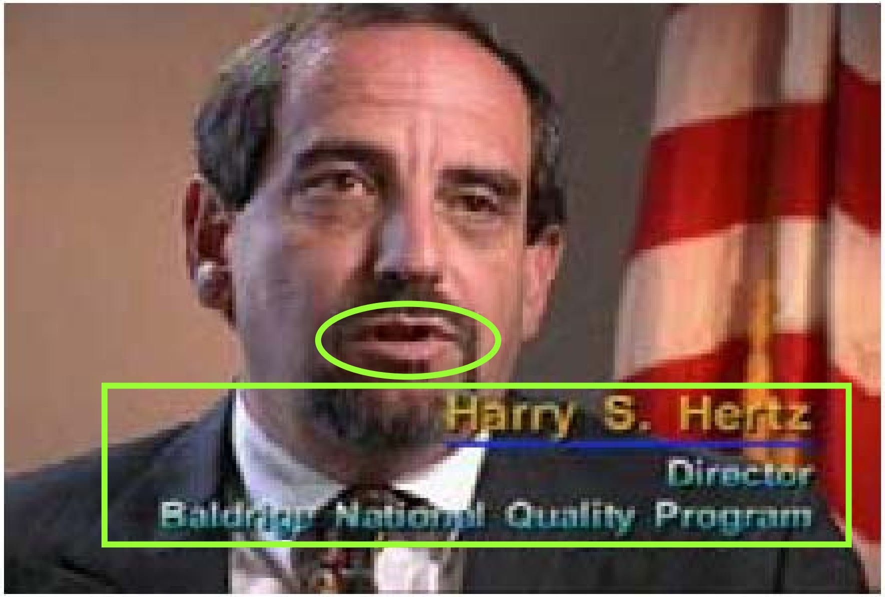
Hints For Better Recognition

- Goal: improve the estimation p(word|acoustic_sig)
- Main idea:

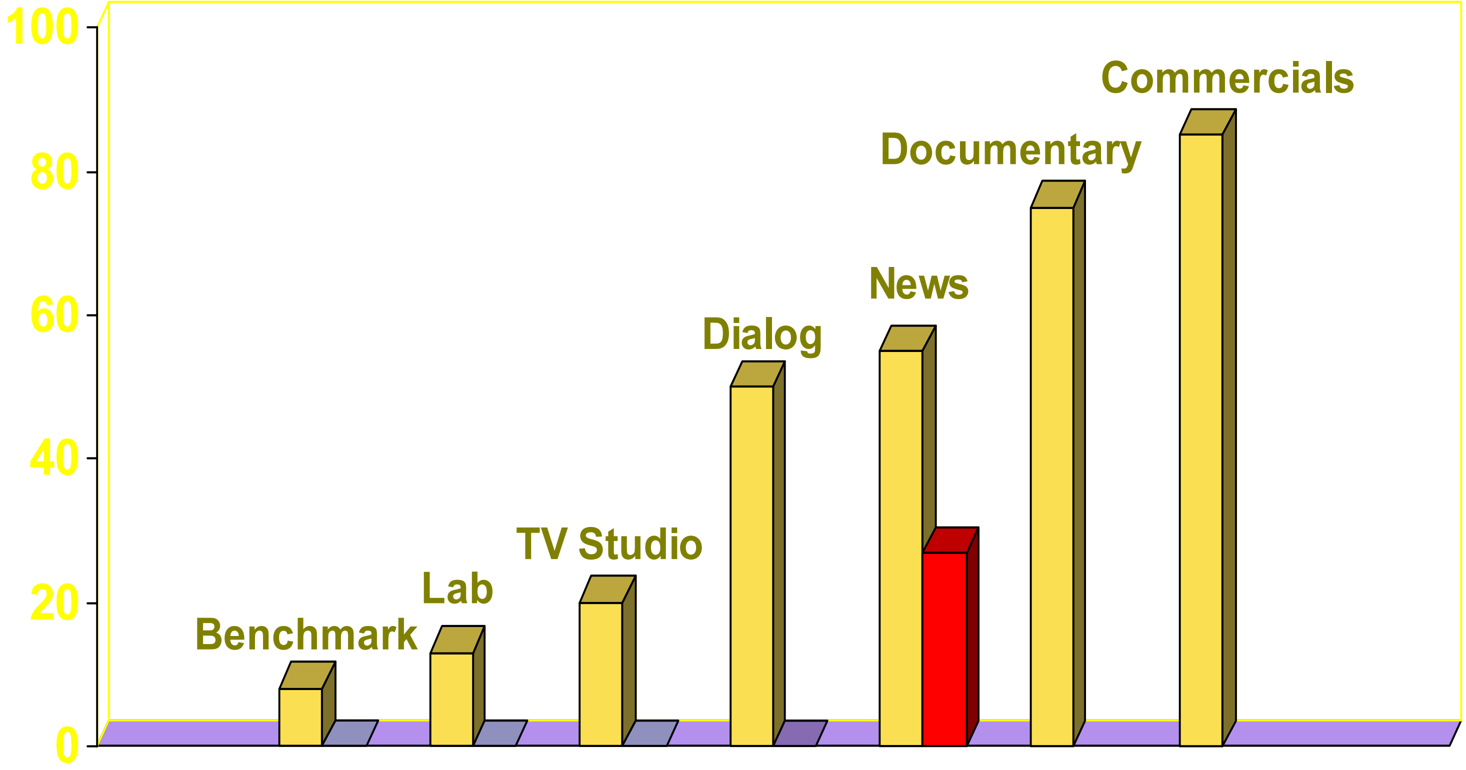
p(word|acoustic_sign) → p(word|acoustic_signal, X)

What could be X?

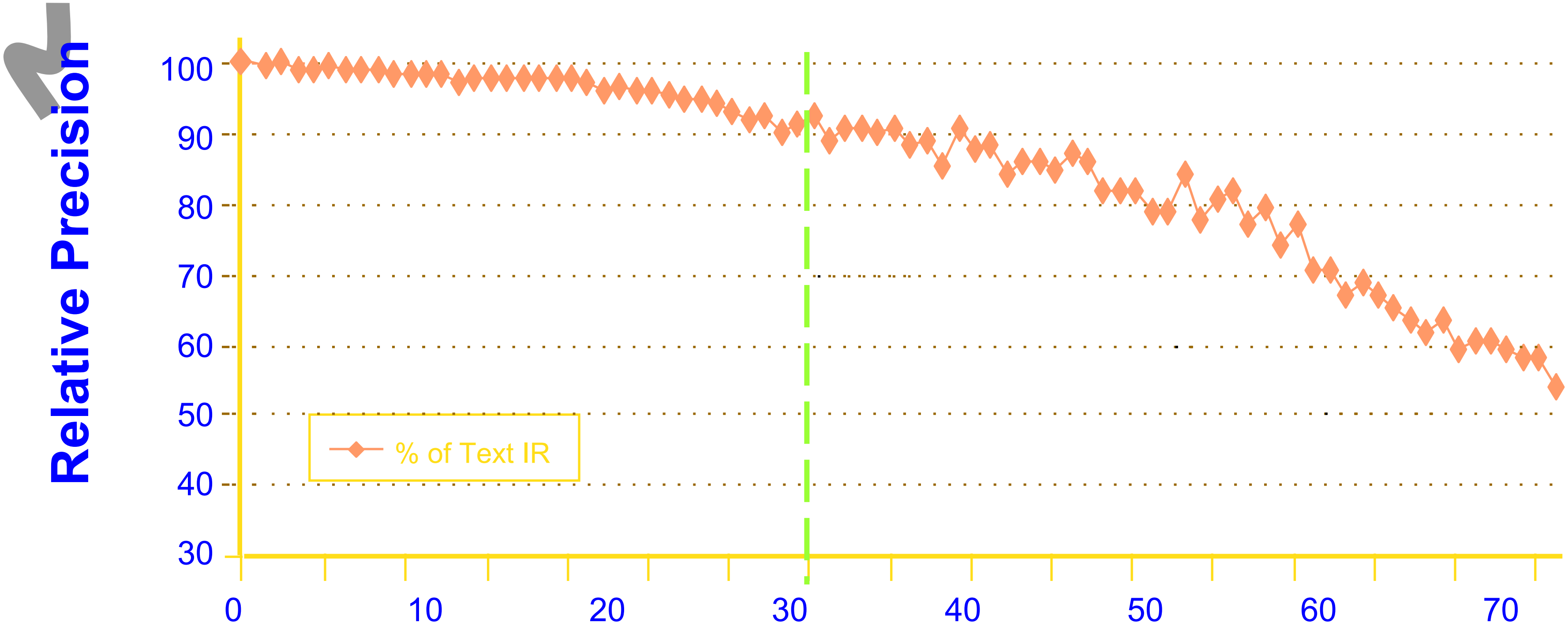
- Topical information
- News of the day
- Image information
  - Lip reading
  - Video Optical Characte Recognition (VOCR)



Speech Recognition Accuracy
Word Error Rate



Information Retrieval Precision vs. Speech Accuracy



A rather small degradation in retrieval when word error rate is small than 30%

Indexing and Search of Multimodal Information, Hauptmann, A., Wactlar, H.
Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP-97), Munich, Germany, April 1997.

Spoken Document Retrieval: Document Expansion

- Motivation: documents are erroneous (or ambiguous)
- Goal: apply expansion techniques to correct the word errors in documents
- Similar to query expansion
  - Treat each speech document as a query
  - Find clear documents that are relevant to speech documents
  - Expand each speech document with the words that are common in the clear documents that are relevant.

Demos

http://images.google.com/

http://video.google.com/

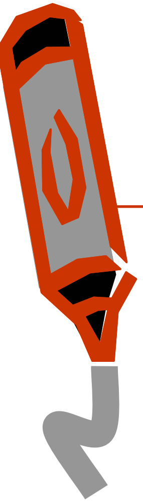
http://www.hermitagemuseum.org/fcgi-bin/db2www/qbicSearch.mac/qbic?selLang=English

http://amazon.ece.utexas.edu/~qasim/research.htm

http://mp7.watson.ibm.com/

http://viper.unige.ch/research/video/

END



exam topics

- Evaluation
  - Recall, precision, E, F
  - AP
  - R-prec
  - PCutoff
  - precision-recall curves
- Retrieval models
  - Boolean
  - Vector space
  - Language modeling
  - Inference networks
- Indexing
  - Manual vs. automatic
  - Tokens, stopping, stemming,
- File organization
  - Bitmaps
  - Signature files
  - Inverted files
- Statistics of text
  - Zipf's law
  - Heap's Law
  - Information theory

- Compression
  - Hufmman
  - LempelZiv
- Relevance feedback
  - Real
  - Assumed
- Clustering
  - Graph, partitioning, nearest neighbor
  - clustering algorithms
- Markov chains
  - stationary distribution
- Page Rank formula
- Metasearch
  - CombSUM
  - Borda
  - Condorcet
- Collaborative filtering
- P2P
  - 3 generations