

T-61.6030 Special Course in Computer and Information Science III: Multimedia Retrieval

Introduction

Markus Koskela

January 18, 2008

Overview of the course

Lecturers	D.Sc. (Tech) Markus Koskela, docent Jorma Laaksonen
Assistant	M.Sc. (Tech) Mats Sjöberg
Credits (ECTS)	5
Semester	Spring 2008 (periods III-IV)
Sessions	On Fridays 12-14 in lecture hall T5
Language	English
Web	http://www.cis.hut.fi/Opinnot/T-61.6030/
Registration	WebTOPI or markus.koskela@tkk.fi

The topic

The **analysis and retrieval** of **multimedia data**.

- combines multiple disciplines:
 - information retrieval
 - computer vision, signal processing
 - pattern recognition, machine learning
 - human-computer interaction
- a combination of image, video, audio, textual data

From

hornet - Google Image Search

http://images.google.com/images?hl=en&q=hornet&btnG=Search+Images&gbv=2

Web Images Maps News Shopping Gmail more Sign in

Google hornet Search Images Search the Web Advanced Image Search Preferences

Moderate SafeSearch is on

Images Showing: All image sizes Results 1 - 20 of about 1,350,000 for hornet [definition]. (0.12 seconds)

<p>asian hornet 440 x 352 - 37k - jpg uneasysilence.com</p>	<p>Hornet 400 x 314 - 49k - jpg www.english-country-garden.com</p>	<p>Hornet Country SpiritWear 266 x 300 - 65k - jpg www.hometcountryspiritwear.com</p>	<p>Above: Bald-Faced Hornet Queen ... 600 x 422 - 74k - jpg www.uos.harvard.edu</p>	<p>Dave Wall's hornet - the ruler has ... 435 x 364 - 41k - jpg naturenet.net</p>
<p>Photo: Close-up of a hornet on a ... 470 x 324 - 30k - jpg animals.nationalgeographic.com</p>	<p>Hornet (Vespa crabro) Czech Republic 600 x 446 - 97k - jpg www.naturephoto-cz.com</p>	<p>Hornet Hornet 2048 x 1536 - 350k - jpg www.wallwork.me.uk</p>	<p>Baldfaced hornet. 504 x 336 - 64k - jpg www.ipm.iastate.edu</p>	<p>Will new hornets emerge from eggs ... 500 x 375 - 317k - jpg www.ipm.iastate.edu</p>
<p>... native European queen hornet, ...</p>	<p>Wasp and Hornet Control 2075 x 1454 - 1507k - ipp</p>	<p>This is a bald-face hornet nest sent ...</p>	<p>The European hornet is a stout ... 727 x 489 - 29k - ipp</p>	<p>More About the Hornet 400 x 305 - 26k - ipp</p>

To

The screenshot displays the IBM Multimedia Analysis and Retrieval System interface. The browser address bar shows the URL: `http://mp7.watson.ibm.com/cgi-imars-bin/WebCBRO6.exe` with the search term `ibm marvel`. The interface includes navigation tabs for Home, Concepts, Clusters, Metadata, Random, and Help. The 'Random' tab is active, showing 'Random Retrieval' results. The top navigation bar includes 'Saved Results: View | Aggregate | Clear'. Below this, there are controls for 'Cols: 5', 'Hits: 50', 'Icons: Lo-Res Hi-Res', 'Text: Hide Show', 'Labels: Hide Show', 'Advanced: Hide Show', and 'Results 1 to 50 of 600'. There are also dropdowns for 'Group by: Shots (ungrouped)', 'Combination: None', 'Aggregation: Avg', and 'Similarity: Color (spatial)'. A search bar with 'Search' and 'Clear' buttons is present. The main content area displays a grid of 20 image thumbnails, each with a numerical ID and a '[More like this]' link. The thumbnails include various scenes: a ship, a financial chart for 'DOWN JONES', a person in a red shirt, a pool table, a group of people, a woman's face, a store interior, a woman at a computer, a logo with the word 'Just', a woman's face, a red textured background, a group of people, a man with a child, a red flower, a chat window with text like 'SydneyAWeb: Hi, you', 'DevinJacks: Want to', 'SydneyAWeb: Sure.', 'DevinJacks: 9:00?', 'SydneyAWeb: Sure.', a group of people, hands holding a green object, a man in a suit, a man in a suit, a person in a white shirt, and a 'NO TRESPASSING SCHOOL DISTRICT PROPERTY' sign.

Requirements

- seminar presentation
- project work (small-scale experiments & written report)
- active participation (one absence is allowed)

Material



Multimedia Retrieval by Blanken, H.M., de Vries, A.P., Blok, H.E., and Feng, L. (Eds.). Springer 2007, 372 p. ISBN: 978-3-540-72894-8.

- + **Machine Learning for Multimedia Content Analysis** by Gong, Y. and Xu, W.. Springer 2007, 277 p. ISBN: 978-0-387-69938-7.
- + **A Unified Framework for Video Summarization, Browsing & Retrieval** by Xiong, Z., Radhakrishnan, R., Divakaran, A., Rui, Y., and Huang T. Elsevier 2006, 296 p. ISBN: 978-0-12-369387-7.
- + relevant journal articles, conference papers

Schedule

Date	Topic	Material
25.1.	Introduction	MR: Ch. 1
1.2.	Metadata, pattern recognition	MR: Ch. 2-3
8.2.	Text retrieval	MR: Ch. 4
15.2.	Image processing	MR: Ch. 5
22.2.	Generative probabilistic models	MR: Ch. 6
29.2.	Speech indexing	MR: Ch. 7
14.3.	Semantic video indexing	MR: Ch. 8
28.3.	Video highlight extraction	UF*: Ch. 3
4.4.	Video retrieval I	MR: Ch. 9
11.4.	Video retrieval II	MR: Ch. 10
18.4.	Interaction	MR: Ch. 11
25.4.	System evaluation	MR: Ch. 13

*: Xiong et al: A Unified Framework for Video Summarization, Browsing & Retrieval



Chapter 1: Introduction

Example usage scenarios

- Journalism, advertising
 - stock photo & video archives
 - “Pretty girl doing something active, sporty in a summery setting, beach - not wearing lycra, exercise clothes - more relaxed in a tee-shirt. Feature is about deodorant so girl should look active - not sweaty but happy, healthy, carefree - nothing too posed or set up - nice and natural looking”
- Digital TV & movies
- Web search
- Medical images
- Surveillance

Getty Images - Search

http://www.gettyimages.com/Search/Search.aspx?query=z.i.H4sIAAAAAAAAEAO29B2AcSZYIj9

Go to the previous version of Editorial site

gettyimages® Sign in | Register Contact us | My account | My cart | International (US)

Home Creative Editorial Footage Music Services

New search Search within Search

matti vanhanen

Advanced Search | Catalyst

Creative images

- RM
- RR
- RF
- Photography
- Illustration
- CDs

Editorial images

All Editorial Any date

Footage

- RR
- RF

Color

- Color
- Black & white
- Horizontal
- Vertical

Refine search

Categories

- Editorial images

People

- Specific people
 - Matti Taneli Vanhanen (31)
 - Tony Blair (11)
 - Vladimir Putin (3)
 - Angela Merkel (3)
 - Recep Tayyip Erdogan (3)
 - See all...
- Gender
- Ethnicity
- Age
- Number of people

Events

- 21 Dec 2007
 - UNS: EU Leaders Hail Expansion Of Passport-Free Schengen Zone (8)

" matti vanhanen " (939 Images found)

FirstLook images only | All search results (939)

Change layout Page 1 of 16 Go to page: Go

<p>Editorial 15 Jan 2008</p> <p>#79007851</p> <p>Finnish Prime Minister Matti Vanhanen sp</p> <p>By: NICHOLAS KAMM</p>	<p>Editorial 15 Jan 2008</p> <p>#79007843</p> <p>Finnish Prime Minister Matti Vanhanen ar</p> <p>By: NICHOLAS KAMM</p>	<p>Editorial 15 Jan 2008</p> <p>#79007827</p> <p>Finnish Prime Minister Matti Vanhanen sp</p> <p>By: NICHOLAS KAMM</p>	<p>Editorial 15 Jan 2008</p> <p>#79007816</p> <p>Finnish Prime Minister Matti Vanhanen sp</p> <p>By: NICHOLAS KAMM</p>

The screenshot shows a web browser window with the URL `http://areena.yle.fi/hae?keyword=nightwish`. The page title is "Hae | Yle Areena | Yle.fi". The YLE Areena logo is at the top left, with the tagline "- Katsele ja kuuntele YLEn ohjelmia". A navigation bar contains links for "Etusivu", "Ohjelmat A-Ö", "Suorat lähetykset", "Podcastit", "Ohjeet", and "Soittolista (0)".

Hakutulokset

Voit tehdä hakutulokseesi erilaisia rajoituksia kielen ja median perusteella. Jos esim. haluat hakea vain suomenkielisiä tv-ohjelmia, rajaa kielivalinnaksi suomi ja mediavalinnaksi Vain videot. Laajimman haun voit tehdä, jos mediavalintana käyttää **Videot ja audiot**, kielivalintana **Kaikki**. Jos ohjelmia ei löydy, tarkista hakurajauksesi.

Hae Areenasta

[Uutiset](#) | [Urheilu](#) | [Ajankohtaisohjelmat](#) | [Asiaohjelmat](#) | [Oppiminen ja tiede](#) | [Kulttuuri](#) | [Viihde](#) | [Musiikki](#) | [Draama](#) | [Lapset](#)

Hakusana

- [Kaikki ohjelmat](#)
- [Kaikki podcastit](#)

Löytyi 2 ohjelmaa. Ohjelmat aikajärjestyksessä. Listauksen esitystapa

Mediarajaukset: « Edelliset 1 Seuraavat »



Tallennetyyppi:

- ▶ **Videot ja audiot**
[Vain videot](#)
[Vain audiot](#)

Kieli:

- ▶ **Kaikki**
[Suomi](#)
[Ruotsi](#)

Ohjelmat

 YleX Tänään Sivunkääntötempussa Tuomas Holopainen 11.01.2008 klo 00:00 fi S	 YleX Aamu Indican Jonsu YleX Aamussa 29.11.2007 klo 12:00 fi S
--	---

Multimedia data

- Originally refers to a collection of media types, nowadays used also for (non-textual) single media types
- Often requires a lot of space for storage, time for processing, bandwidth for transfer
- Effective presentation not trivial either
 - indexing
 - summarization

Metadata: descriptive

- Descriptive data in some specified format:
author, date, format, size, ...
- Dublin Core, MPEG-7

Internet Archive: Details: Night of the Living Dead

http://www.archive.org/details/night_of_the_living_dead

Home Animation & Cartoons | Arts & Music | Computers & Technology | Cultural & Academic Films | Ephemeral Films | Movies | News & Public Affairs | Non-English Videos | Open Source Movies | Prelinger Archives | Sports Videos | Video Games | Vlogs | Youth Media

Search: Movies [Advanced Search](#) Anonymous User ([login](#) or [join us](#))

View movie



[View thumbnails](#)
Run time: 1:35:17

Stream ([help](#))

[64Kb MPEG4](#) (dialup)
[256Kb MPEG4](#) (broadband)

Play / Download ([help](#))

[64Kb MPEG4](#) (108 MB)
[256Kb MPEG4](#) (246 MB)
[MPEG1](#) (918 MB)
[MPEG2](#) (4.1 GB)

All files: [FTP](#) [HTTP](#)

NO RIGHTS RESERVED


Resources

[Bookmark](#)
[Report errors](#)

Night of the Living Dead (1968) **George A. Romero**



Click to play video

[embedding and help](#)

In this classic yet still creepy horror film, strangers hold up in a rural Pennsylvania farmhouse and battle constant attacks from dead locals who have been brought back to life by mysterious radiation.

Note: This item contains a user-contributed srt subtitle file. To use this file you must download an srt compatible player and point it at the correct video and srt files (google for srt subtitles). We include this file for advanced users who may wish to use it, however the Archive does not support any player that displays subtitles stored external to the video they are intended to be used with, nor can we vouch for the quality or completeness of the subtitling effort.

WARNING - This is a graphic horror film with some gory special effects.

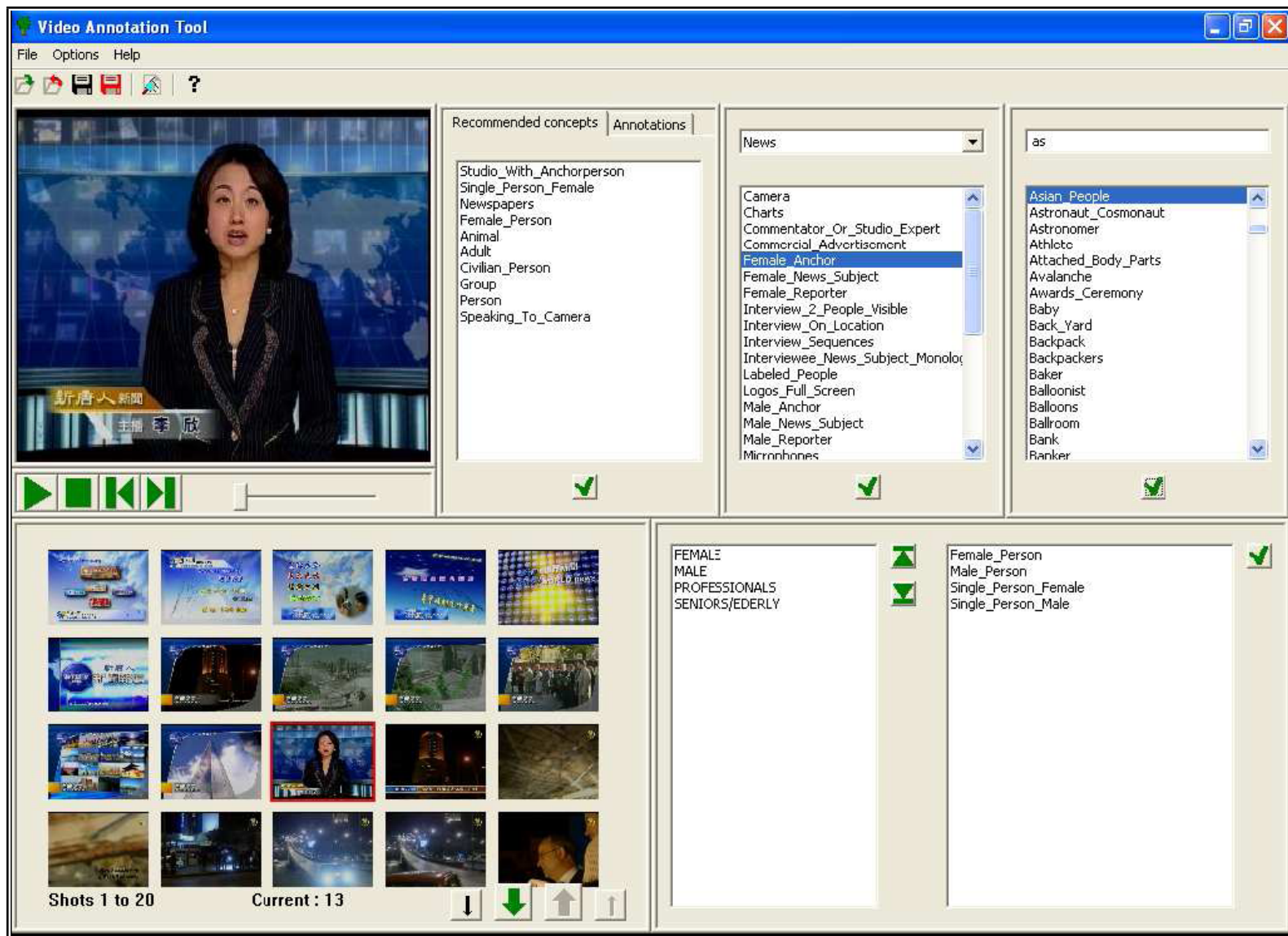
You can find more information regarding this film on [its IMDb page](#) .

This item is part of the collection: [Feature Films](#)

Director: George A. Romero
Producer: Karl Hardman, Russell Streiner
Audio/Visual: sound, b&w
Keywords: [horror](#)

Metadata: annotations

- Lists of keywords, free-format, ontology-based
- Manual annotation
 - time consuming, expensive, subjective
 - flickr, LabelMe, Google Image Labeler, ...
- Automatic annotation
 - a popular research topic

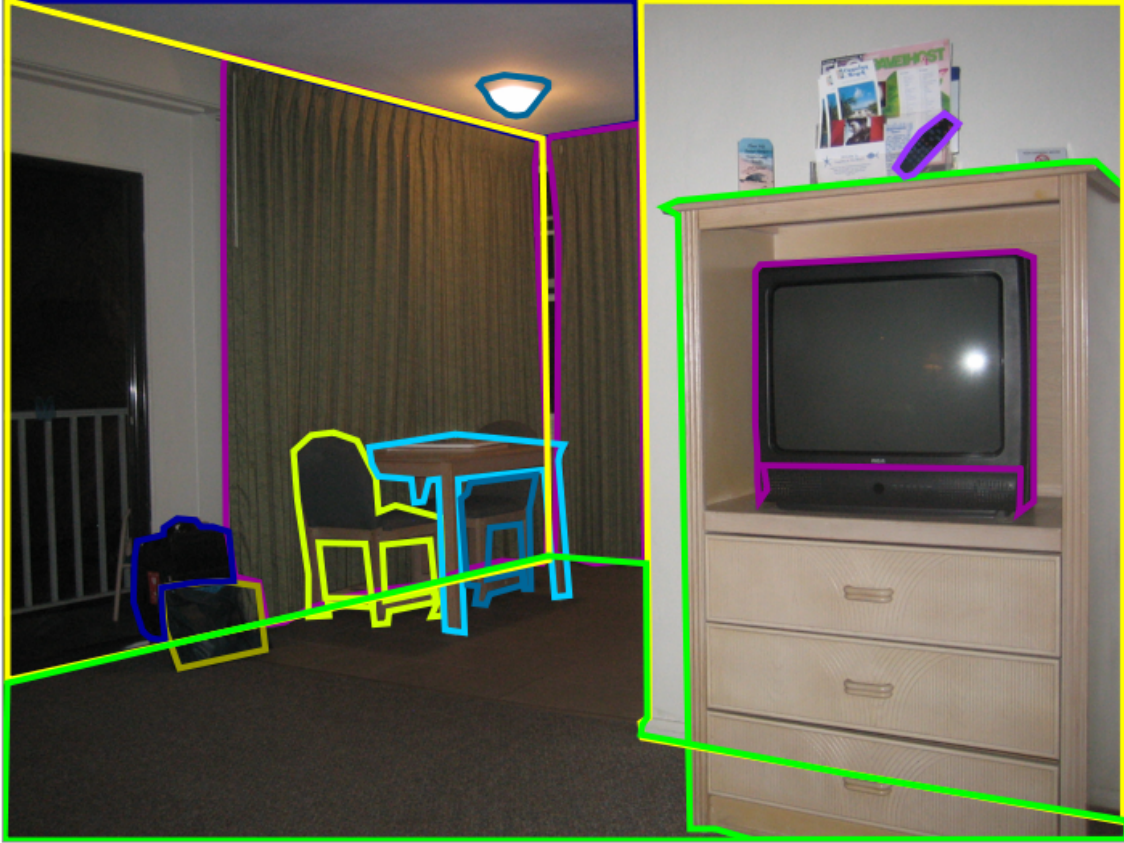


LabelMe: The open annotation tool

http://labelme.csail.mit.edu/tool.html?collection=LabelMe&mode=i&folder=static_hotel_room_indoor&im

LabelMe Please [contact us](#) if you find any bugs or have any suggestions.

Label as many objects and regions as you can in this image [Show me another image](#)




Sign in (why?)


There are **207222** labelled objects

Instructions (Get more help)

Use your mouse to click around the boundary of some objects in this image. You will then be asked to enter the name of the object (examples: car, window).

Good  Bad 

Labeling tools

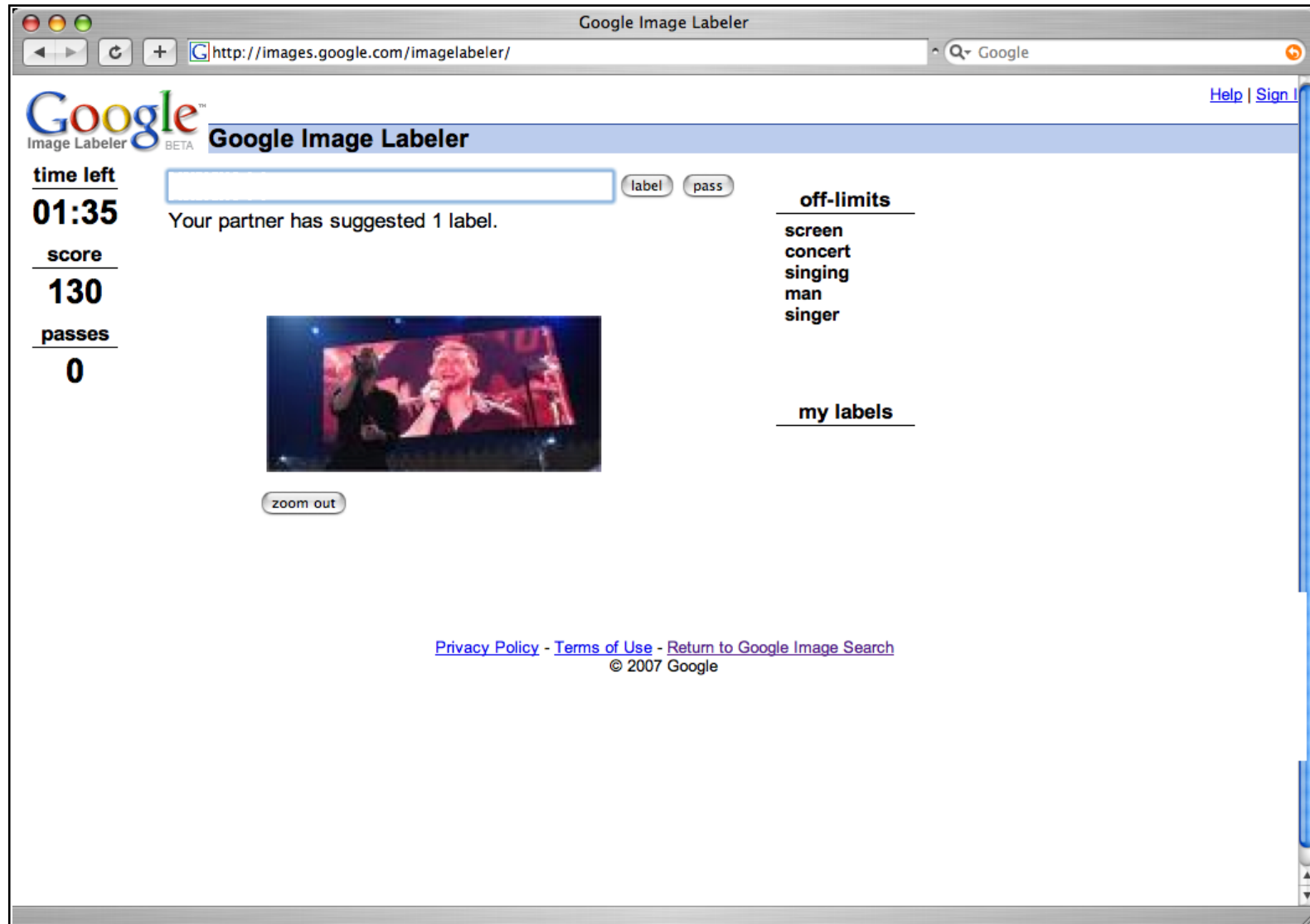
  

Erase segment **Zoom** **Fit Image**

Polygons in this image (XML)

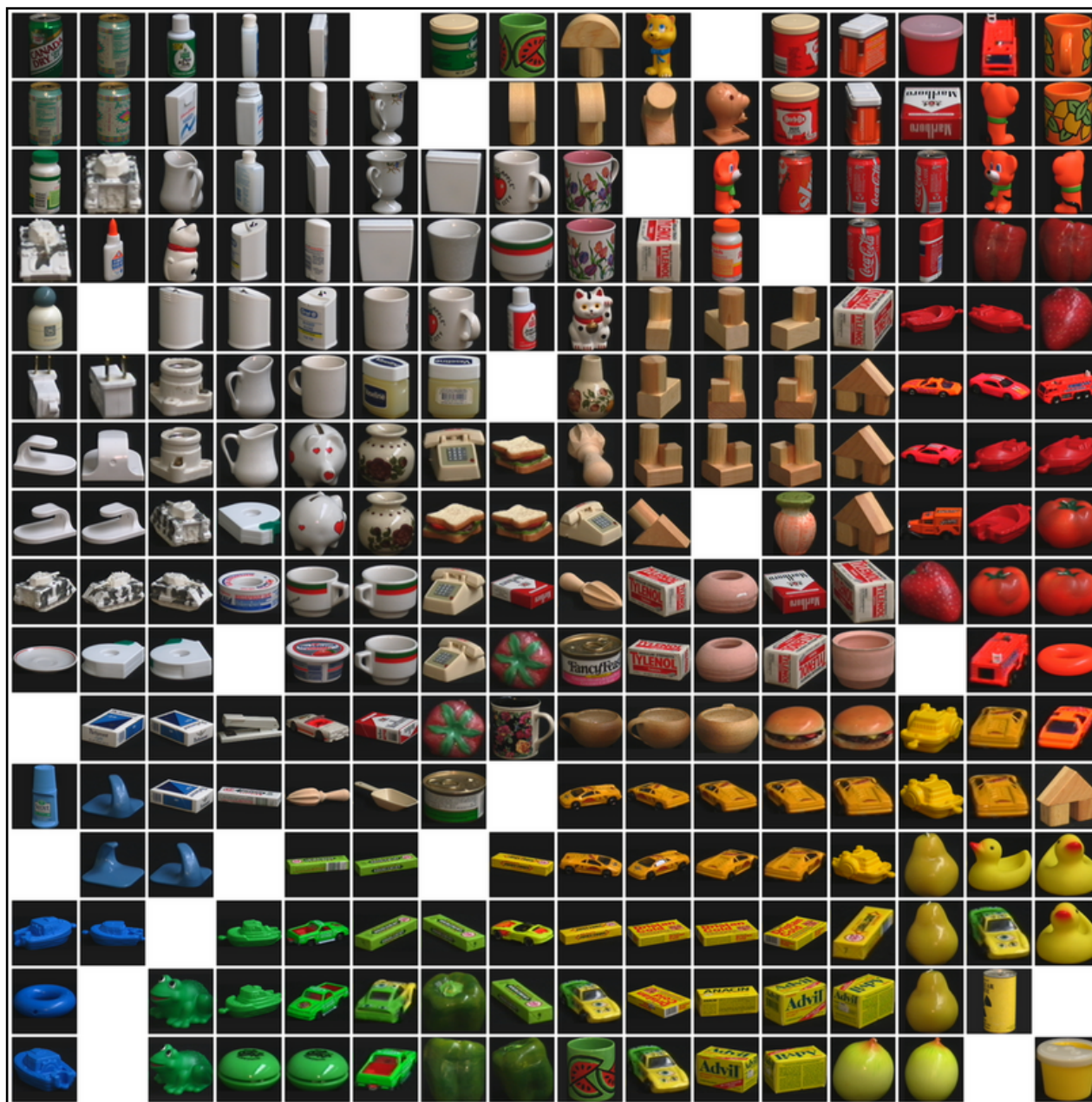
- [tv](#)
- [ceiling](#)
- [curtain](#)
- [curtain](#)
- [furniture](#)
- [bag](#)
- [shoolbag](#)
- [chair](#)
- [table occluded](#)
- [chair occluded](#)

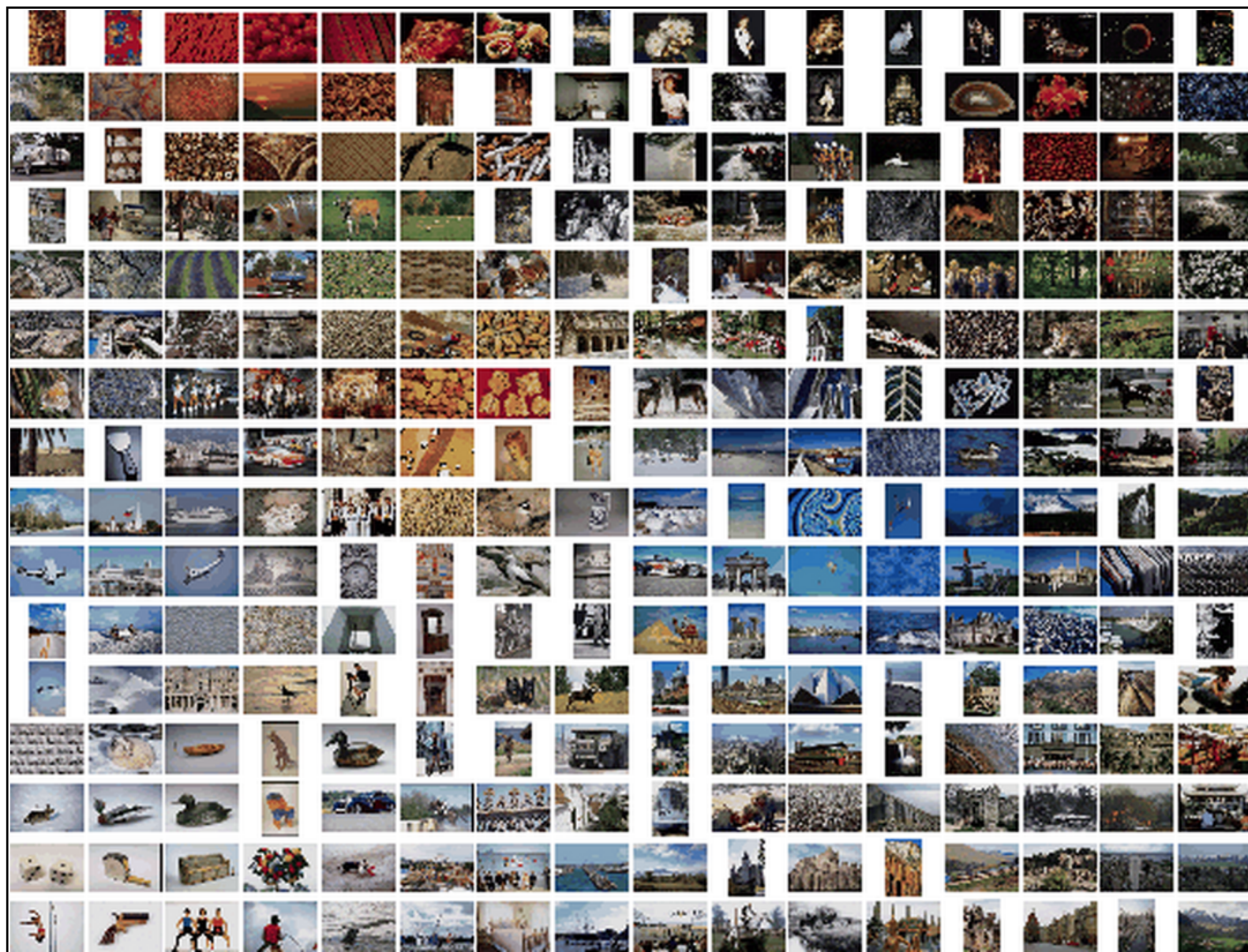
Done



Metadata: low-level features

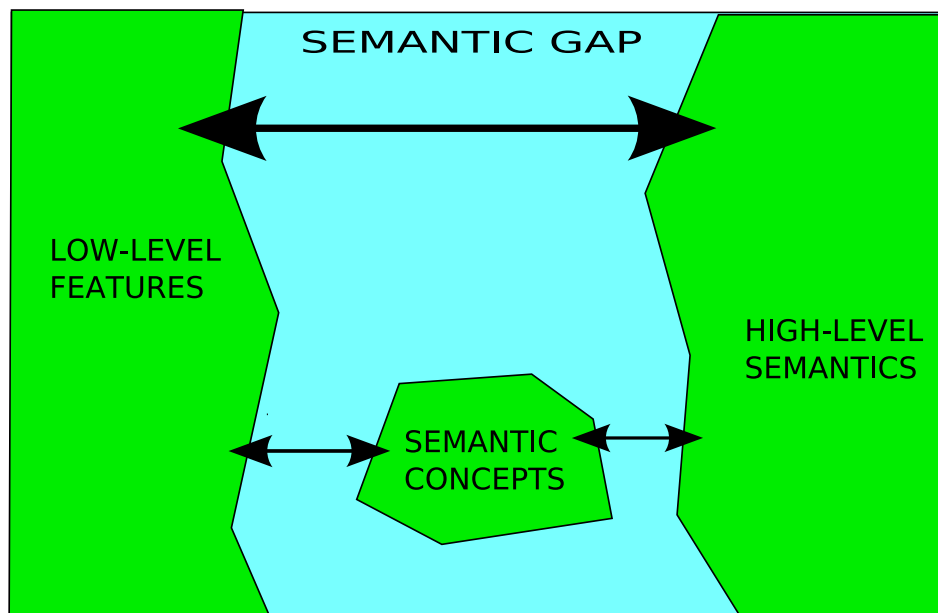
- Features that can be extracted automatically, usually represented as fixed-dimensional vectors
- **Text**: bag of words, n-grams, ...
- **Audio**: ZCR, DFT, silence ratio, ...
- **Images**: color histogram, color layout, edge histogram, Gabor texture, shape features, ...
- **Video**: all of above, motion, ...





Metadata: high-level features

- End users are interested in semantically meaningful metadata, not low-level features: **the semantic gap**



- An important question: **How to derive high-level features from the low-level ones?**



Multimedia queries

- More complex than text searches or SQL queries
- **Degree of similarity** instead of exact match
- Different types of queries:
 - query by descriptive metadata
 - query by annotations
 - query by low-level and high-level features
 - query by example
 - query by user profile

IBM Multimedia Analysis and Retrieval System (IBM T. J. Watson Research Center)

http://mp7.watson.ibm.com/cgi-imars-bin/WebCBR06.exe?SessiD=707608733&LIST=25576&H ~ ibm marvel

Home Concepts Clusters Metadata Random Help

Visual search with IDs "25682" (77817) Saved Results: View | Aggregate | Clear

Cols: 5 Hits: 50 Icons: Lo-Res Hi-Res Text: Hide Show Labels: Hide Show Advanced: Hide Show





















Group by: Shots (ungrouped) Combination: None Aggregation: Avg Results 1 to 50 of 77817

Similarity: Color (spatial) Page 1

LIKE@25682 Search Clear

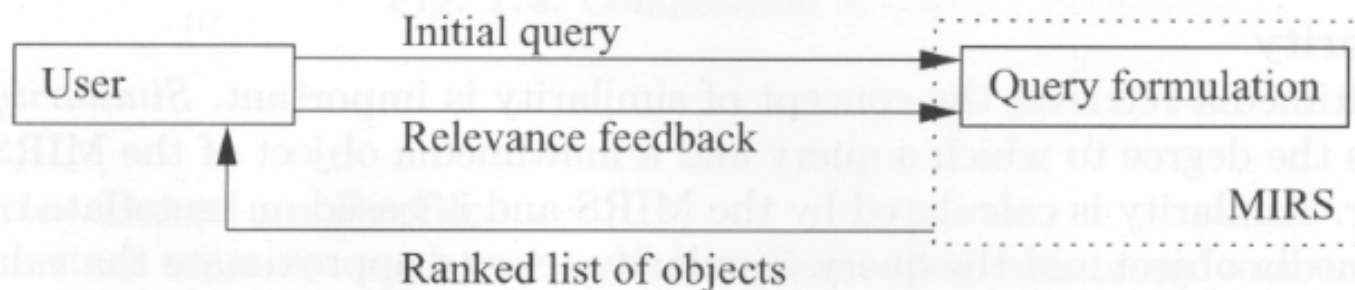
Refine with: FACE | MILITARY | ANIMAL | FLAG-US | PEOPLE-MARCHING

Filter by VIDEOS: TRECVID2005 73 (16) | TRECVID2005 76 (6) | TRECVID2005 107 (6) | TRECVID2005 48 (5) | TRECVID2005 18 (3)

 25682. More like this	 [0.848] More like this	 [0.846] More like this	 [0.838] More like this	 [0.835] More like this
 [0.829] More like this	 [0.829] More like this	 [0.829] More like this	 [0.828] More like this	 [0.824] More like this
 [0.823] More like this	 [0.809] More like this	 [0.801] More like this	 [0.797] More like this	 [0.795] More like this
 [0.794] More like this	 [0.794] More like this	 [0.793] More like this	 [0.793] More like this	 [0.793] More like this

Relevance feedback









- Formulating a multimedia query can be difficult
- One solution: iterative and interactive refinement of the original query
- The user evaluates the relevance of retrieved items, which is then fed back to the system











































PicSOM

http://www.cis.hut.fi/picsom/ftp.sunet.se/Q:080116:153458:423680:180aaa!S=sphinx.hut.fi

Selected objects:

Query objects:

Continue query Change settings Restart Restart query

Browsing

- Another solution: let the user to quickly scan through the (summarized) data
- Requires an effective interface
- Browsing and retrieval can be combined

Video Retrieval Tool - j89ur

File Task View Tools Help

View selection Task definition Task time left

Q B R

Browser settings

Group similar shots in:

- columns
- videos

Similar match timeline:

- visually
- conceptually
- textually

Browsing history

Timeline of video no. 189 (20041105_230000_NBC_NBCPHILA23_ENG.mpg)

 are we going to be seeing that kind of wind	 the latest victim of a wardrobe malfunctioned	 you hit themal this weekend this story you	 now by stick out of the bill leigh valley the	 got to cover of the air by friday night live buys	 MONDAY 11PM	
---	---	--	---	---	-----------------	--

Shots similar to timeline

 had a golden opportunity to never			 new jersey will have a new governor in just	 in her election carries focus is clear there is is	 MONDAY 11PM	
	 everything should be posted	 at least 15 civilians in the beijing	 tested awarded stranded to test your	 but she shot to do the best you can and i'm	 MONDAY 11PM	 reality is
				 we had a really good phone call	 MONDAY 11PM	
 dampor on our playoff additional friday night		 nineteen ninety eight thousand	 is in trouble with the law tonight's swimmer	 cherish and capital covered capital in this	 postal service	
 and a dividend of one of its didn't realize that the	 and without grain our ninety nines radio		 eagles are sitting pretty at seven and one b.	 to the white house this was an historic election	 text not available	 have lost her baby she's four months pregnant

