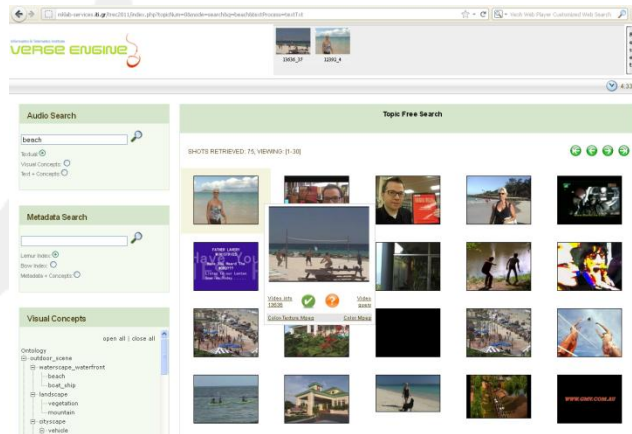


ITI-CERTH @ Known Item Interactive Search Task



Stefanos Vrochidis

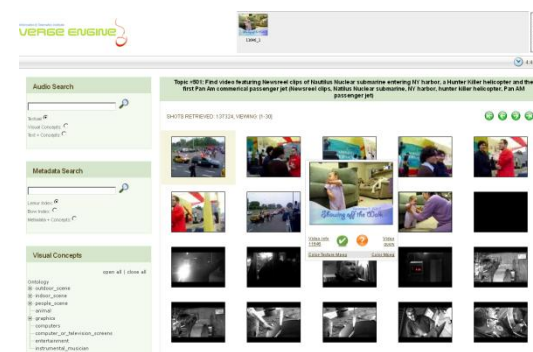
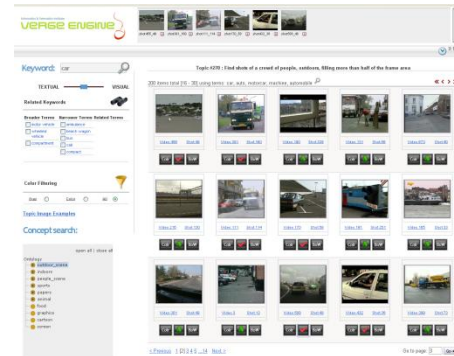
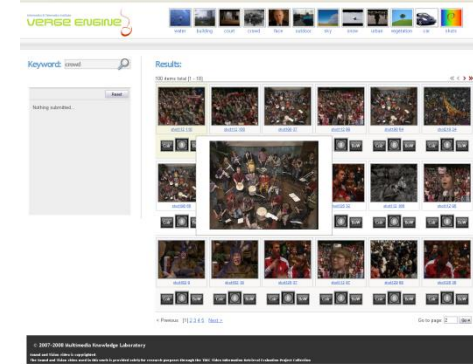


Informatics and Telematics Institute
Centre for Research and Technology Hellas

A Mourtzidou, P. Sidiropoulos, S. Vrochidis, N. Gkalelis, S. Nikolopoulos, V. Mezaris, I. Kompatsiaris, I. Patras, "ITI-CERTH participation to TRECVID 2011", TRECVID 2011 Workshop, December 2011, Gaithersburg, MD, USA.

ITI-CERTH @ TRECVID

- Search Task
 - TRECVID 2006-2008
 - Under COST 292
 - TRECVID 2009-2010
- Instance Search Task
 - TRECVID 2010
- Known Item Search Task
 - TRECVID 2010-2011
- VERGE Video Search Engine
 - Interactive Video Search



Problem Description

- **Known Item Search Task**
 - The user is supposed to know a video in advance
 - A detailed textual video description is provided
 - Time for search is limited to 5 minutes
- **Interactive Search - Ideas**
 - The system needs to respond fast
 - Fusion could assist in combining efficiently results
 - Could we exploit the implicit user feedback?
 - Take into account the semantic relations of metadata



VERGE

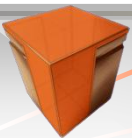
- Interactive Platform
- Web-based
- Technologies
 - Apache
 - PHP
 - Javascript
 - Lemur
- Modules
 - Metadata Search (Lemur)
 - ASR Search (Lemur)
 - Visual concept search
 - Visual Similarity search
 - Fusion
 - PLSA based search
- URL
 - <http://mklab-services.iti.gr/trec2011>

The screenshot displays the VERGE web interface with the following components:

- Metadata Search (text, bow and fusion search)**: A search bar with a dropdown menu.
- Audio Search (text, concept and fusion search)**: A search bar with a dropdown menu.
- Visual Concepts**: A list of concepts including outdoor_scene, watercape_waterfront, landscape, cityscape, weather, snow, sky, daytime_outdoor, nighttime, indoor_scene, people_scene, animal, graphics, computers, and computer_tv_television_screens.
- History**: A list of recent searches including "reversed moon - made huffd" and "helixlike moon - made lemur".
- Main Results Area**: A grid of video thumbnails with a "RETRIEVED: 24, VIEWING: [1, 24]" indicator.
- Stored Results**: A section on the right side of the results area.
- Video Shots & Side Shots**: A section at the bottom left of the results area.
- Metadata Bow**: A section at the bottom right of the results area.
- Mpeg-7 Color- Texture Search**: A section at the bottom right of the results area.
- Mpeg7 - Color Search**: A section at the bottom right of the results area.
- Search History**: A section at the bottom left of the results area.
- Visual Concept Search**: A section at the bottom left of the results area.

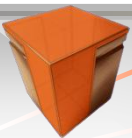
Video Indexing

- Temporal Indexing
 - Shot Segmentation
 - Representative keyframe extraction
- Visual similarity Indexing
 - MPEG-7
- Textual Data Indexing
 - ASR
 - Metadata
 - Lemur
- Visual concepts extraction
 - Results from the SIN task
 - SURF descriptors
 - Video tomographs



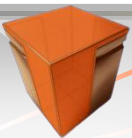
Implicit User Feedback

- User actions are recorded during search sessions
- Mouse hover time on presented shots was measured
- Concept Fusion
 - Attention Fusion Method
- ASR and Concept Fusion
 - Attention Fusion Method
 - SVM regression model (after enough examples)
- Metadata and Concept Fusion
 - At video level
 - Attention Fusion Method
 - SVM regression model (after enough examples)



Semantic Relatedness

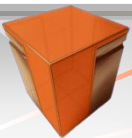
- Indexing using the semantic relatedness of metadata
 - Metadata
 - Bag of Words approach
 - Vector with 1000 words
 - Video represented as word count histogram
 - Multiplied with Wordnet distance vector
 - “vector” similarity was used
 - Probabilistic Latent Semantic Analysis
 - 25 latent topics
- Functionalities
 - Video similarity (based on metadata)
 - Metadata search



Experiments

- 4 runs
- Combinations of modules

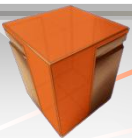
Modules	Run IDs I.A.YES_ITI-CERTH_x			
	x=1	x=2	x=3	x=4
ASR Lemur text	no	yes	yes	yes
ASR fusion	yes	no	no	no
Metadata Lemur text	no	yes	yes	no
Metadata BoW text	no	no	yes	yes
Metadata fusion	yes	no	no	no
High Level Visual concepts	yes	yes	no	yes



Experiment Design

- Participants
 - Gender
 - 6 males
 - 2 females
 - Topic distribution
 - 6 or 7 topics each
 - Education
 - PhD students
 - Research Assistants
 - Short tutorial

TOPICS/RUNS	run1	run2	run3	run4
500				
501				
502	MALE 1	MALE 3	MALE 4	FEMALE 2
503				
504				
505				
506				
507				
508	FEMALE 1	MALE 1	MALE 5	MALE 6
509				
510				
511				
512				
513				
514	MALE 4	FEMALE 1	MALE 2	MALE 5
515				
516				
517				
518				
519				
520				
521	MALE 3	MALE 6	FEMALE 2	MALE 2
522				
523				
524				

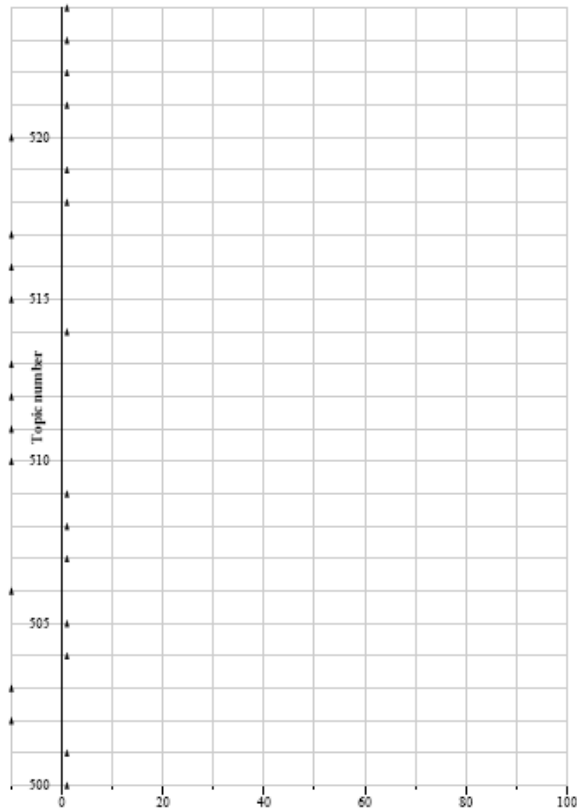


Experiments

Run ID: ITI-CERTH
Processing type: Interactive
System training type: A (only IACC training data)
Condition: YES (IACC.1 *_meta.xml used)
Priority: 1

Across 25 test topics
Mean inverted rank: 0.560
Mean elapsed time (mins): 3.257
Mean user satisfaction (1-7 best): 6.000

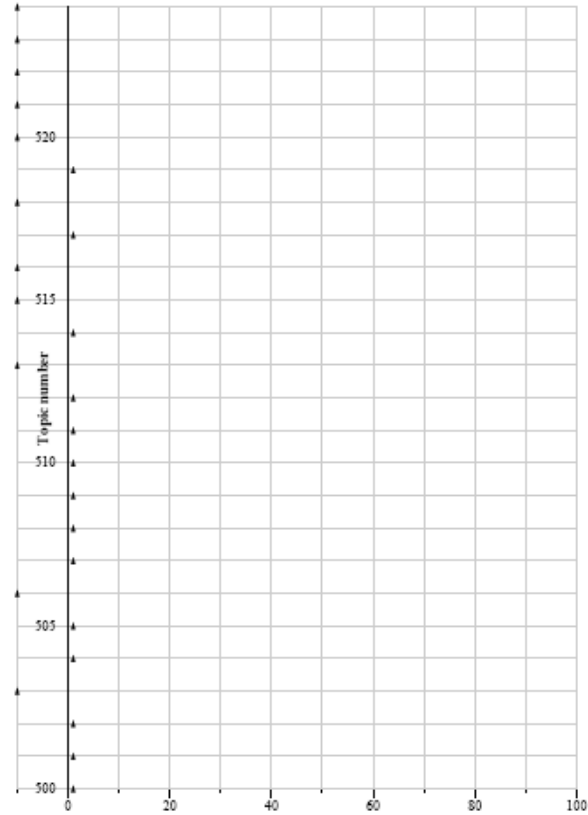
Rank (1--top) of found KI by Topic (Misses graphed left of y-axis)



Run ID: ITI-CERTH
Processing type: Interactive
System training type: A (only IACC training data)
Condition: YES (IACC.1 *_meta.xml used)
Priority: 2

Across 25 test topics
Mean inverted rank: 0.560
Mean elapsed time (mins): 3.284
Mean user satisfaction (1-7 best): 6.000

Rank (1--top) of found KI by Topic (Misses graphed left of y-axis)

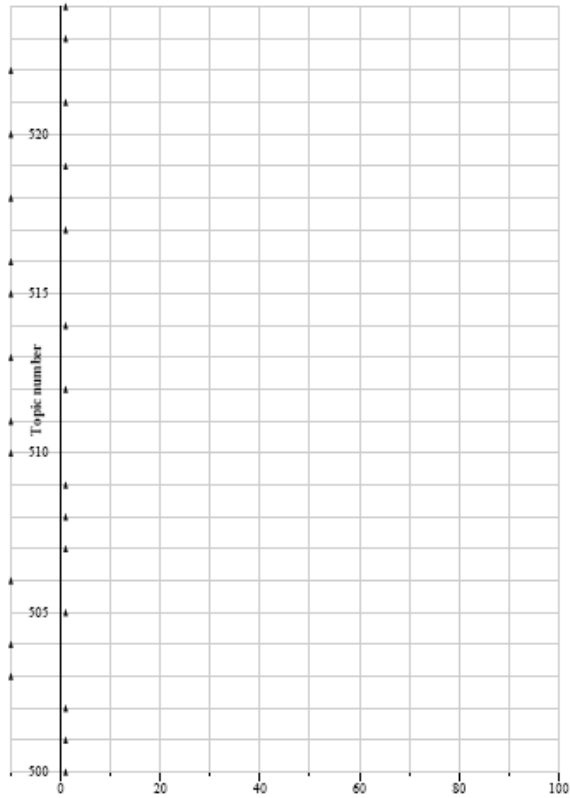


Experiments

Run ID: ITI-CERTH
Processing type: Interactive
System training type: A (only IACC training data)
Condition: YES (IACC.1 *_meta.xml used)
Priority: 3

Across 25 test topics
Mean inverted rank: 0.560
Mean elapsed time (mins): 3.274
Mean user satisfaction (1-7 best): 5.000

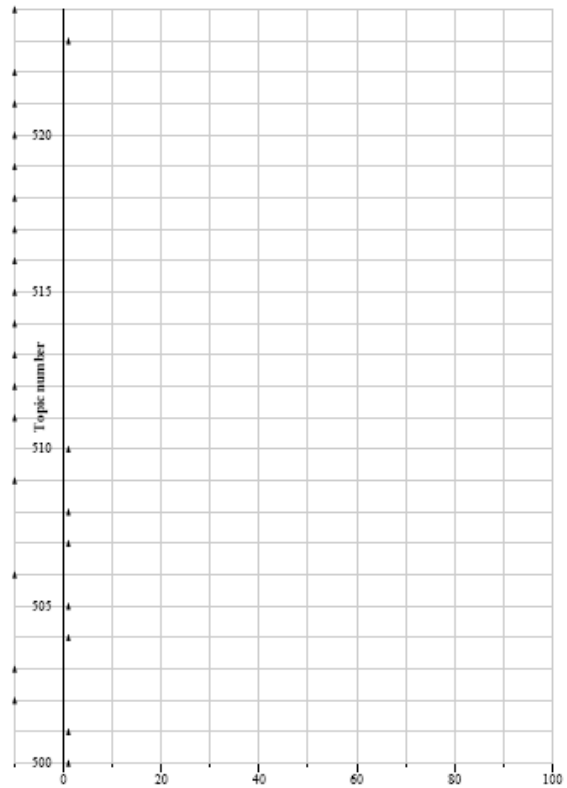
Rank (1=top) of found KI by Topic (Misses graphed left of y-axis)



Run ID: ITI-CERTH
Processing type: Interactive
System training type: A (only IACC training data)
Condition: YES (IACC.1 *_meta.xml used)
Priority: 4

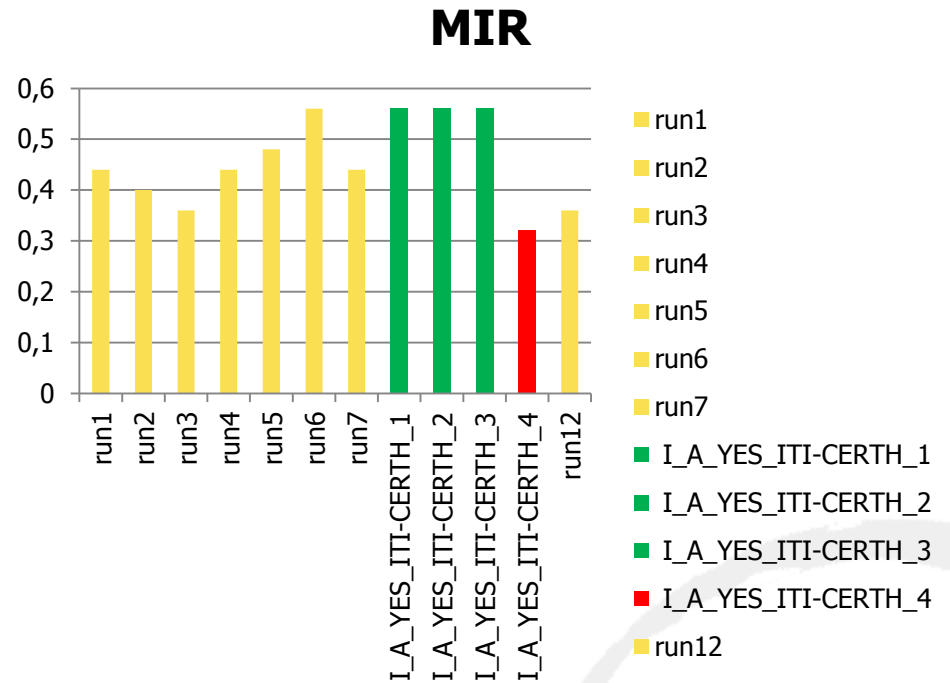
Across 25 test topics
Mean inverted rank: 0.320
Mean elapsed time (mins): 4.072
Mean user satisfaction (1-7 best): 5.000

Rank (1=top) of found KI by Topic (Misses graphed left of y-axis)



Results

Runs and systems	MIR	CORRECT (/25)
run1	0,44	11
run2	0,4	10
run3	0,36	9
run4	0,44	11
run5	0,48	12
run6	0,56	14
run7	0,44	11
I_A_YES_ITI-CERTH_1	0,56	14
I_A_YES_ITI-CERTH_2	0,56	14
I_A_YES_ITI-CERTH_3	0,56	14
I_A_YES_ITI-CERTH_4	0,32	8
run12	0,36	9



Conclusions

- Results

- The most efficient module is still the metadata and ASR search
 - Many modules to use in a limited time
 - Users are still more familiar with simple text search
- Time was limited to see whether implicit feedback could improve the results
- Fusion could be promising in such limited time tasks
- SIN low performance did not affect the system
- Semantic relatedness analysis didn't show any improvement
- Maybe more simple search tasks could be used to evaluate these new functionalities.

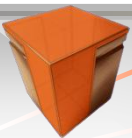
- Task

- Some times the textual topic description doesn't give the right impression for the video
- In many cases knowledge of the topic makes a difference (e.g. Ellis island -> New York, statue of liberty)



Future Work

- Video based preview
- Faster Fusion
- Reduce search options that might confuse the user
- Keep track which specific module produced a correct result
- Query expansion



Thank you!

CERTH-ITI / Multimedia Group

<http://mklab.itigr>

