

# semantic multimedia content analysis



*knowledge in multimedia content  
ist event 2006*

yannis avrithis, ivml-ntua

helsinki, wednesday 22.9.2006

# consider this

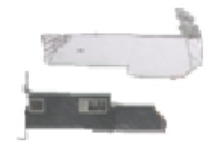


foliage/trees  
color/texture  
hard to merge  
too much variety



house

color/texture/shape  
spatial structure  
(roof over wall)  
hard to separate



person  
specialized detector  
hog/sift/svm  
deformable  
articulated



snow  
color/texture  
easy



air mattress!!  
shape/edges/lines  
unusual / invariants?  
occlusion / out of context!

# use case / domain



news  
video



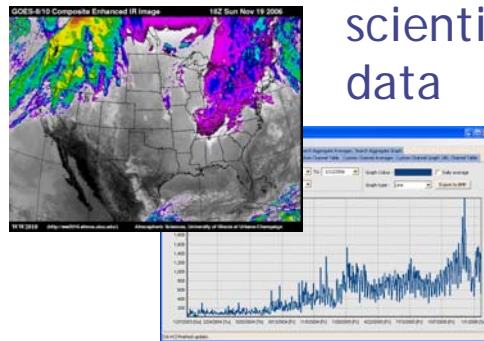
security



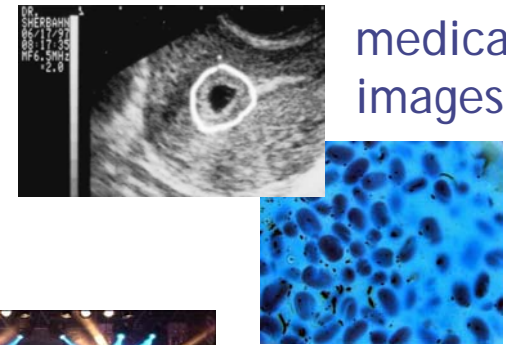
surveillance



personal  
photos



scientific  
data



medical  
images



culture



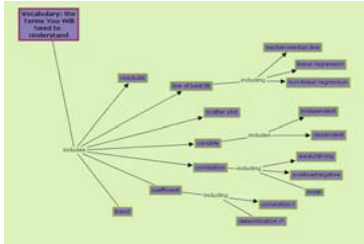
sports



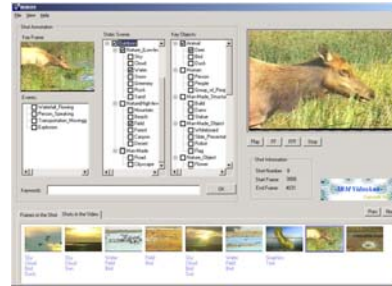
music

# not just analysis!

vocabulary

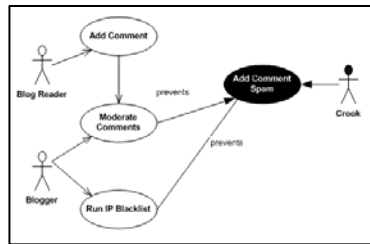


annotation tool



annotation

```
- <ann:depicts>  
- <rdf:Statement rdf:ID=  
- <rdf:subject>  
- <beach:Human rdf:  
<ml:hasName r  
<ml:isStanding>  
</beach:Human>
```



use case



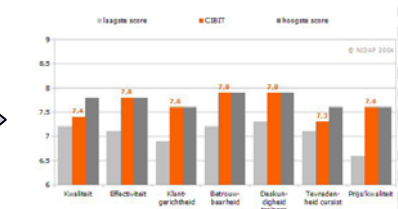
dataset



retrieval tool

36. Explosion\_Fire: S  
37. Natural-Disaster:  
38. Maps: Shots depic  
39. Charts: Shots depi

queries



benchmark

# nist trecvid benchmark



news video data  
~250hr, ~150GB  
multilingual transcripts



people



building



animal



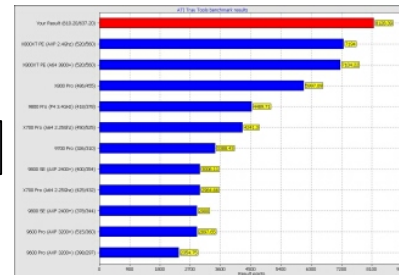
vegetation

shot-level annotation  
~450 concepts (Iscom)

- **visual** features (color, texture, shape, etc.)
- **multimodal** feature combination (audio, text, transcripts)
- supervised learning, **generic** detector -> svm
- use **context** & production style hints

- ~95% p&r, shot cuts
- ~20% p-r breakeven, "feature" detection
- ~42% ap, inter. search
- **speech** -> fixed vocab, limited speakers, us only

evaluation



benchmark over ~80  
participants worldwide



average precision (ap)  
measure: combine  
precision & recall

# annotation on the web



Images Showing: All image sizes



horse show Buckingham.jpg  
1531 x 1263 pixels - 209k - jpg  
e-advancement.csupomona.edu



horse-and-little-horse  
800 x 600 pixels - 57k - jpg  
7art-screensavers.com

0:52  
Time Left

The ESP Game

0250  
score



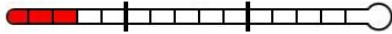
Taboo Words  
FOOTBALL  
GAME

Your Guesses

Type your next guess:

Pass

Your partner has entered a guess



Flag

Adblock



Search

Photos Groups People

Everyone's Photos

bird

Full text Tags only

We found 655,931 photos about bird.

View: Most relevant Most recent Most interesting



King Vulture

Uploaded on 28 September 2005



By True\_Bavaria...

See more photos, or visit his profile.

bird, nature, animal, animals ...

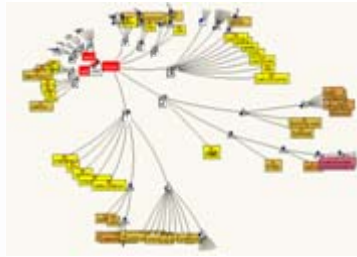


Please contact us if you find any bugs or have any suggestions.

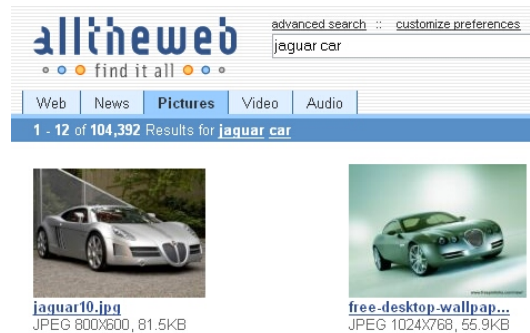
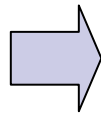
Label as many objects and regions as you can



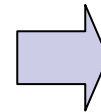
# building image ontologies



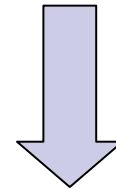
pruned wordnet  
24,000 leaves



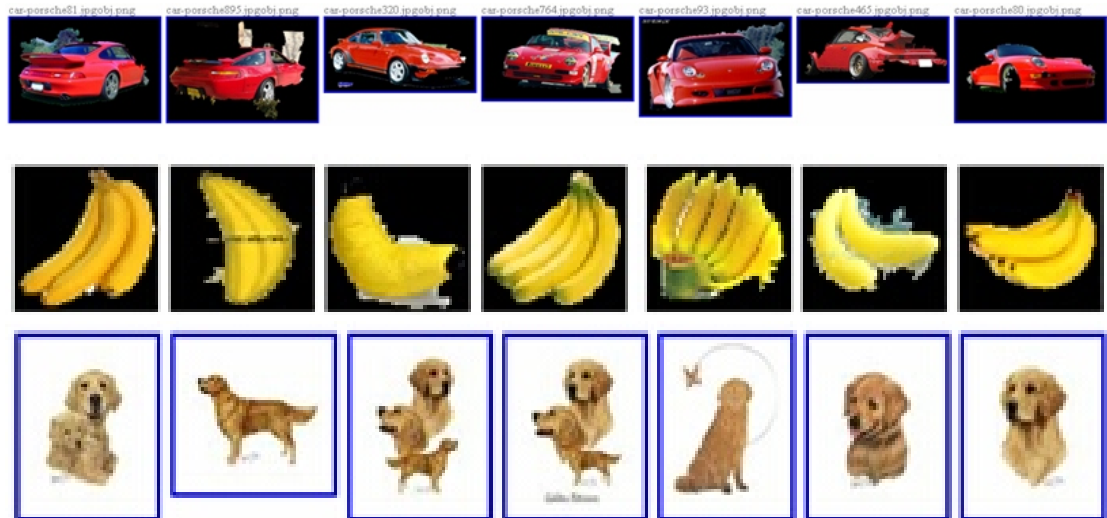
web search  
portrayable object  
+ hypernym



foreground segmentation



clustering  
+ sorting

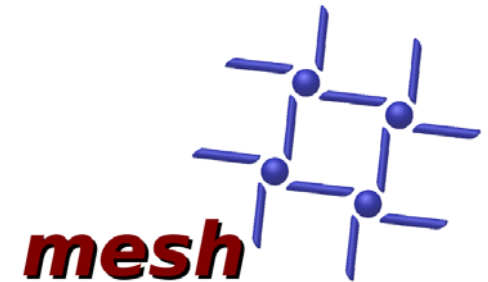
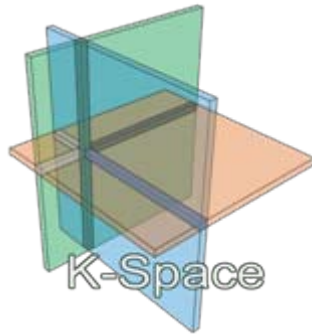


# more points to remember

- multimedia analysis: **starting** to give results
- **huge** research on its own, not just tool as part of projects
- **broad domain** <-> poor performance
- **semi-automatic**: still useful, reduce annotation cost
- **ontologies** useful only when used in reasoning  
-> more expressivity, **rules** ???
- promote european/global **benchmarking**  
(datasets, groundtruth, ontologies, etc.)



# projects



...see you at...



1st international conference on semantic and digital media technologies

december 6-8, 2006 · athens, greece

<http://samt2006.org>

**WIAMIS 2007** 8th International Workshop on Image Analysis for Multimedia Interactive Services  
6-8 June, 2007, Santorini, Greece



W3C Activities

Incubator Activity

- **Multimedia Semantics XG**

[Incubator Activity](#) > W3C Multimedia Semantics Incubator Group



thank you!

[iavr@image.ntua.gr](mailto:iavr@image.ntua.gr)

<http://image.ntua.gr/~iavr>