The Mediamill 2005 semantic video search engine

Marcel Worring, Richard van Balen, Jan-Mark Geusebroek, Jan van Gemert, Bouke Huurnink, Dennis Koelma, Giang Nguyen, Ork de Rooij, Arnold Smeulders, Cees Snoek

Intelligent Systems Lab Amsterdam
University of Amsterdam

TRECVID interactive search

- Task
  - Search the right shots for some given information need

- Data
  - Chinese, Arabic, American News data

- Constraint
  - Per topic time is 15 minutes maximum

Search Topic

1. Condoleeza Rice
2. Iyad Allawi
3. Omar Karami
4. Hu Jintao
5. Tony Blair
6. Mahmoud Abbas
7. graphic map of Iraq, Bagdhad marked
8. two visible tennis players on the court
9. people shaking hands
10. helicopter in flight
11. George W. Bush entering or leaving a vehicle
12. something on fire with flames and smoke
13. people with banners or signs
14. people entering or leaving a building
15. a meeting with a large table and people
16. a ship or boat
17. basketball players on the court
18. one or more palm trees
19. an airplane taking off
20. a road with one or more cars
21. one or more military vehicles
22. a tall building
23. a goal being made in a soccer match
24. office setting

0

24 search topics

Generic/specific
Object/Setting/People

isis

Remember the 101 Concepts

1 - 35

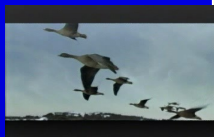Aircraft Allawi Anchor Animal Arrafat Baseball Basketball Beach
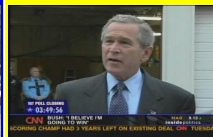
Bicycle Bird Boat Building Bus Bush jr. Bush sr. Candle
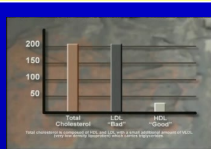
Car Cartoon Chair Chart Clinton Cloud Corp. leader Court
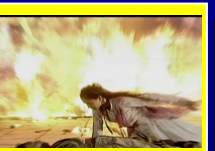
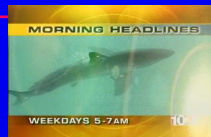Crowd Cycling Desert Dog Drawing Drawing/cartoon Duo-anchor Entertainment

Explosion Face Female

The 101 Concepts

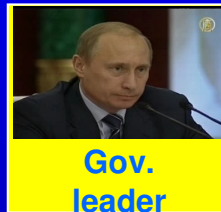| Fire weapon | Fish | **Flag** | Flag USA | Food | Football | Golf | Gov. building |
Gov. leader	Graphics	Grass	Hassan Nasrallah	Horse	Horse racing	House	Hu Jintao
Indoor	Kerry	Lahoud	Male	**Map**	**Meeting**	**Military**	Monologue
Motorbike	**Mountain**	Natural disaster	News paper	Night fire	**Office**	Outdoor	Overlayed text
People	**People marching**	Police / security					

The 101 Concepts

| Powell | Prisoner | Racing | Religious leader | River | **Road** | Screen | Sharon |
Sky	**Smoke**	Snow	**Soccer**	Split screen	Sports	Studio	Swimming pool
Table	**Tank**	**Tennis**	**Tony Blair**	**Tower**	**Tree**	Truck	Urban
Vegetation	**Vehicle**	Violence	**Walking**	**Water body**	Waterfall	Weather news	

Interactive retrieval

- If there is a simple mapping from topic to concepts
  - …… and the concept is of sufficient quality
  - performing a concept query yielding a ranking is sufficient

- One additional observation
  - If one element of a program is relevant, other elements of the program are often also relevant

The Cross Browser
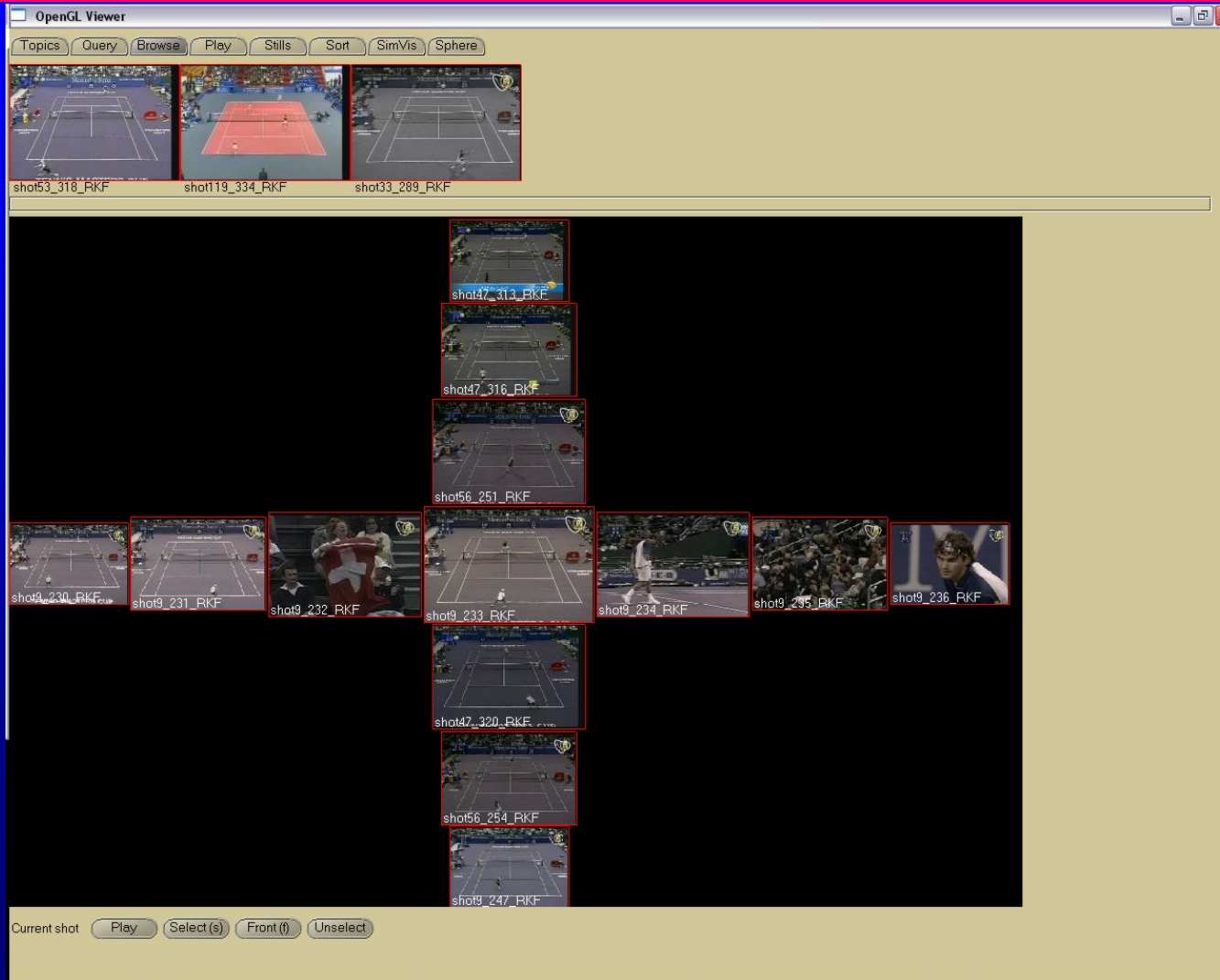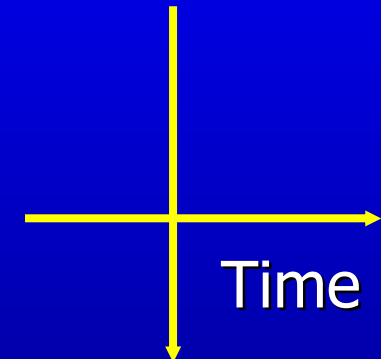
So what if the relation is less obvious

- Use the structure of the multimodal concept space
  - How are different information sources related?
  - Can we use combinations of different concepts?

- But the space is very complicated
  - So can we guide the user through concept space?

Video feed

Shot segmentation
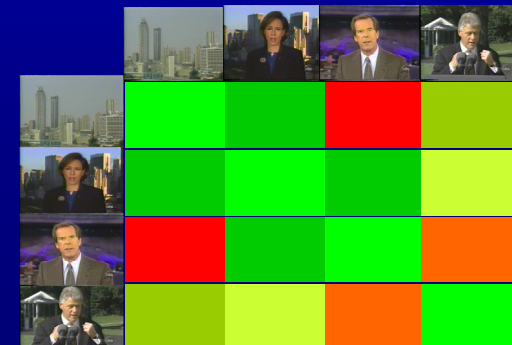
List of shots

Content analysis

Shots with concept vector

Building 0.9
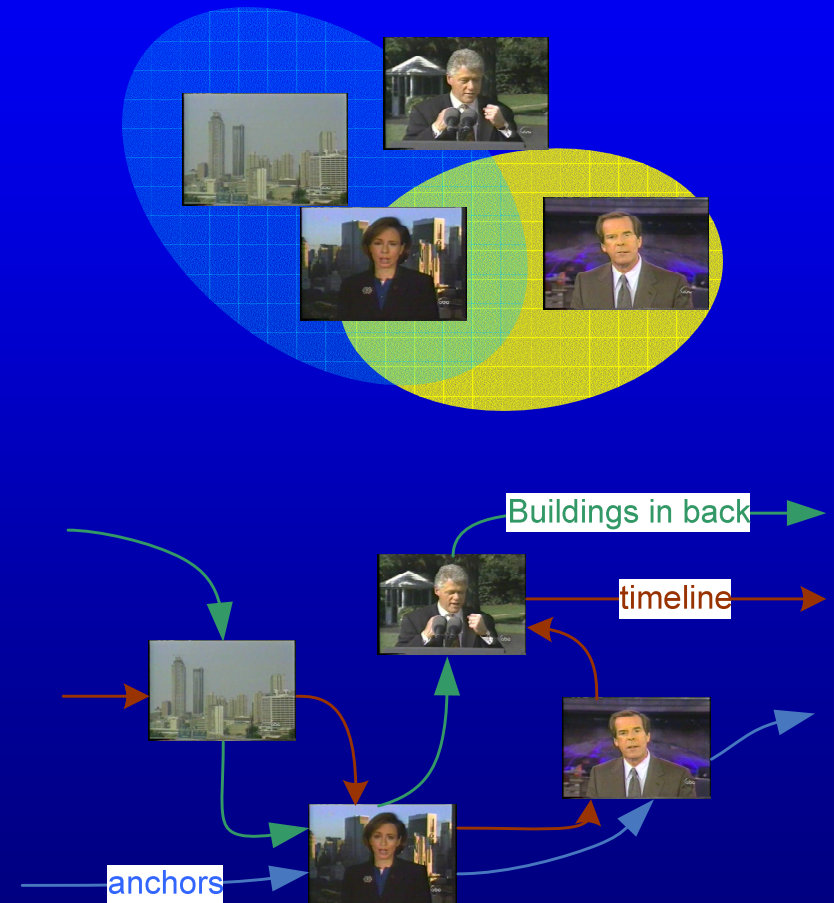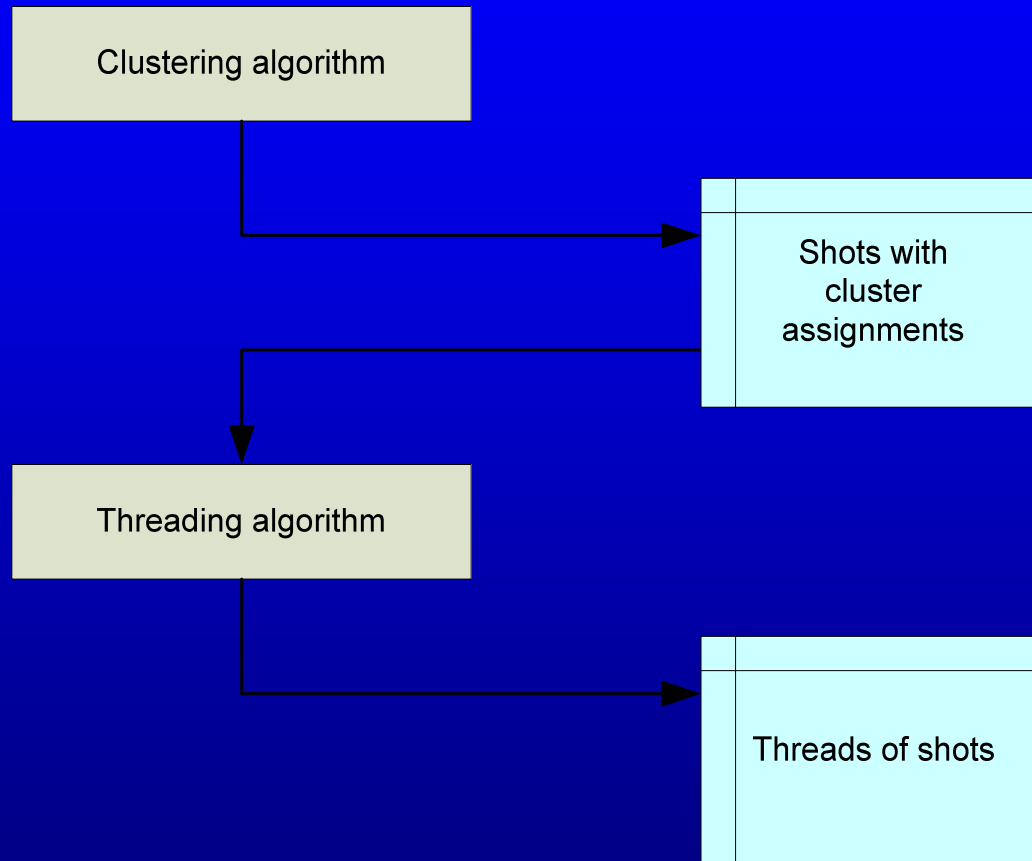
Building 0.6
Anchor 0.8
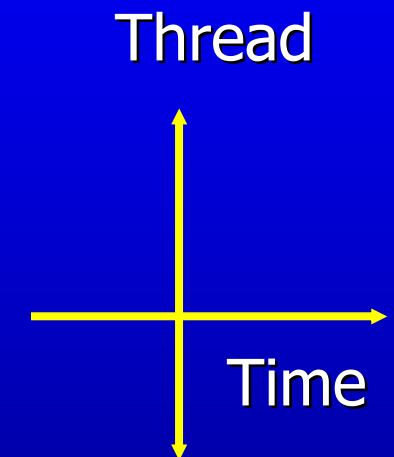
Anchor 0.9

Clinton 0.8
Building 0.4
Vegetation 0.6

Distance function

Shots with distances

Finding semantic threads

Clustering algorithm

Shots with cluster assignments

Threading algorithm

Threads of shots

Buildings in back

timeline

anchors

The Sphere Browser

Thread

Time
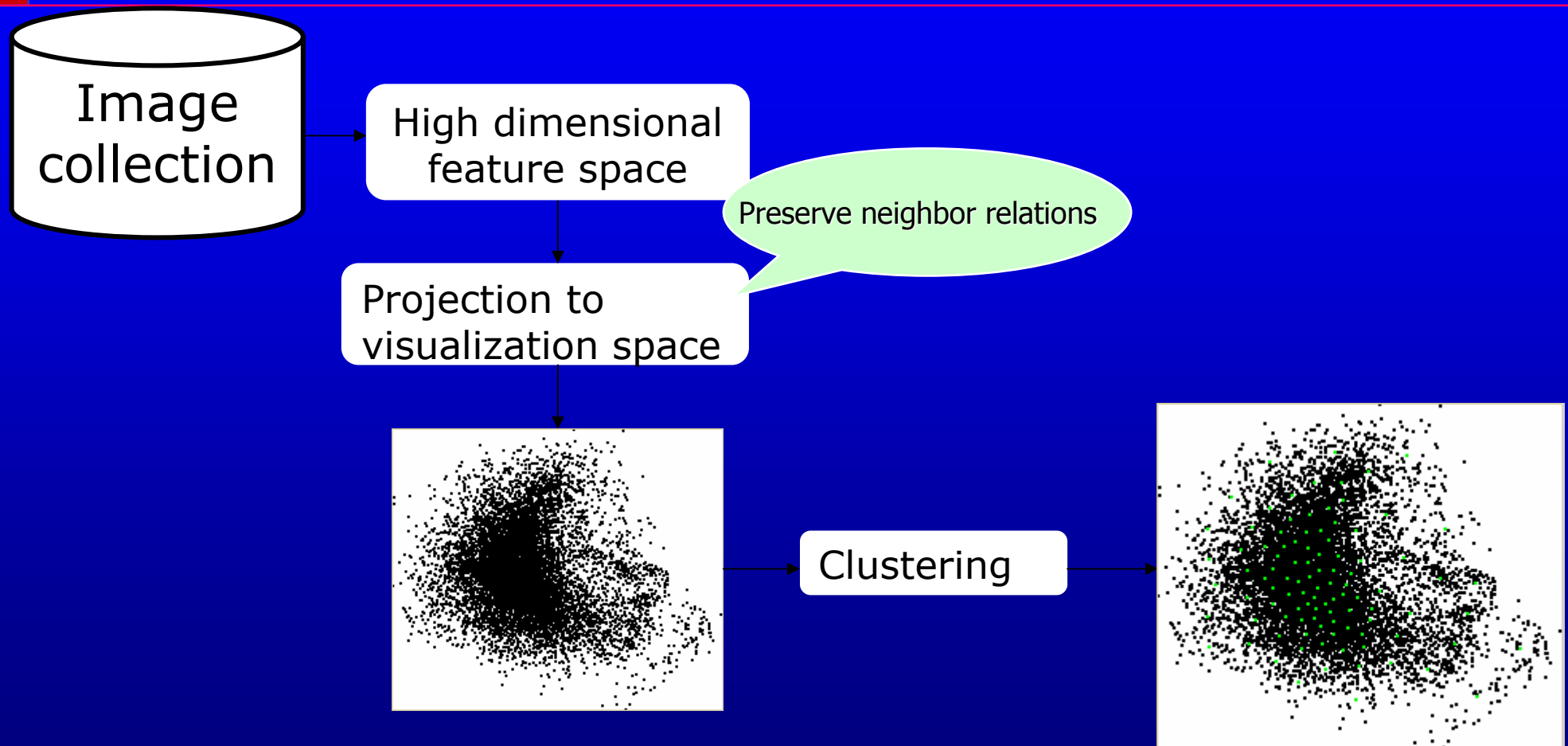
And if there are no concepts?

- Visual browsing
  - Exploring a large part of the whole dataset
  - Make use of clusters in visual appearance

- Also relevant if there are sub-concepts to be found
  - Especially if the instantiations of the sub concepts are visually similar

Video feed

Shot segmentation

List of shots

Content analysis

Shots with concept vector

Visual feature vector

Distance function

Shots with distances
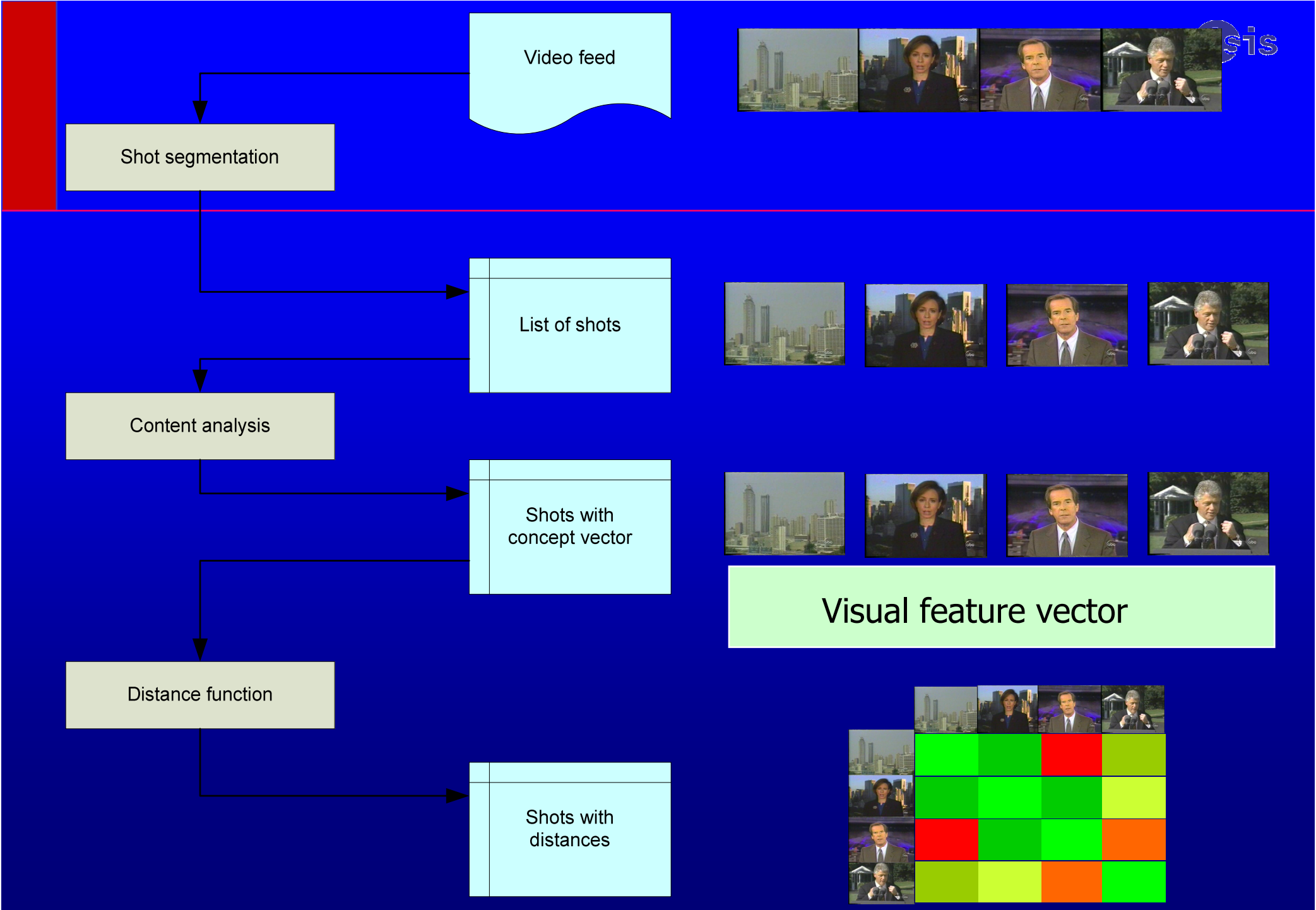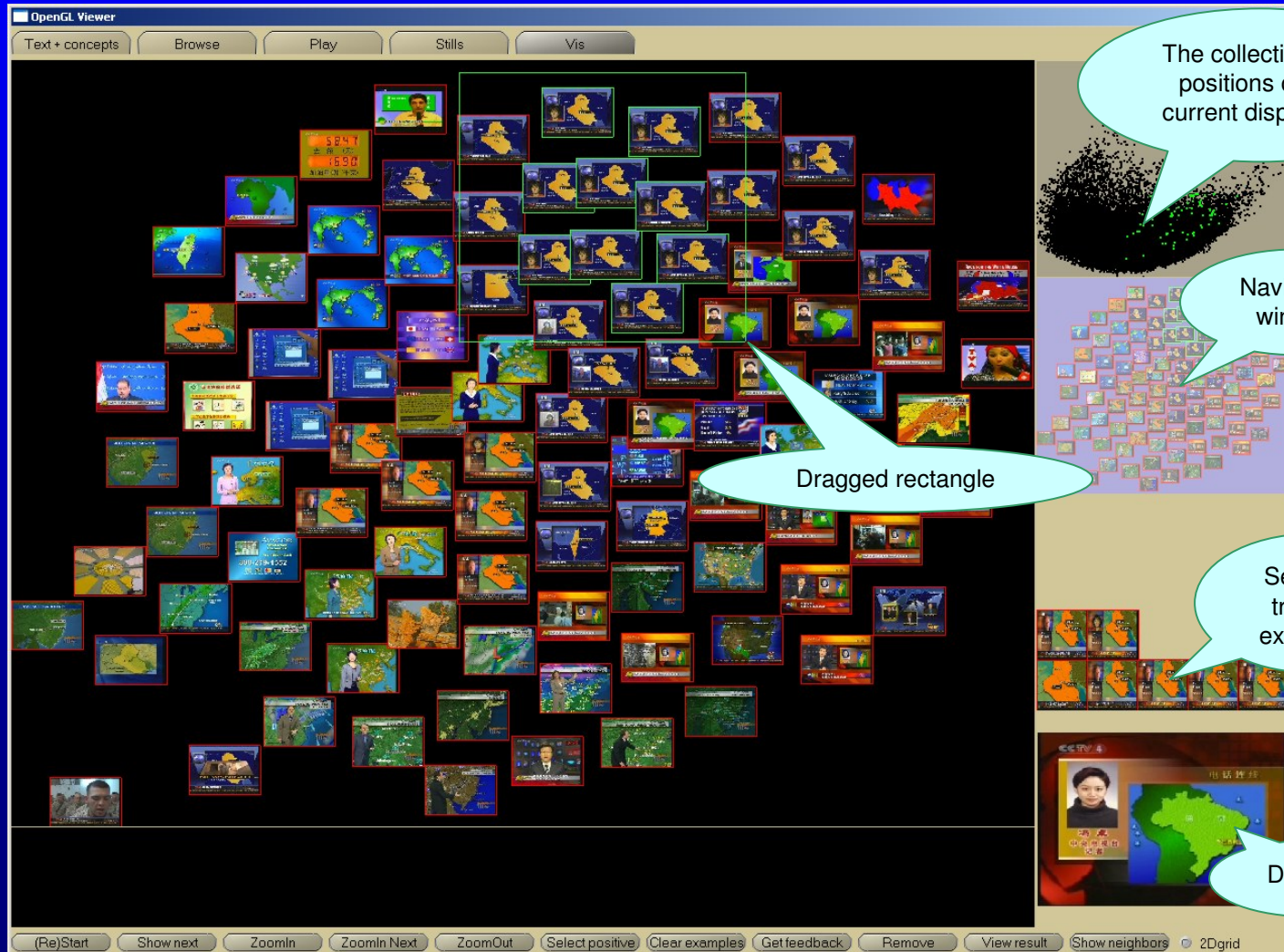
The offline stage

On-line
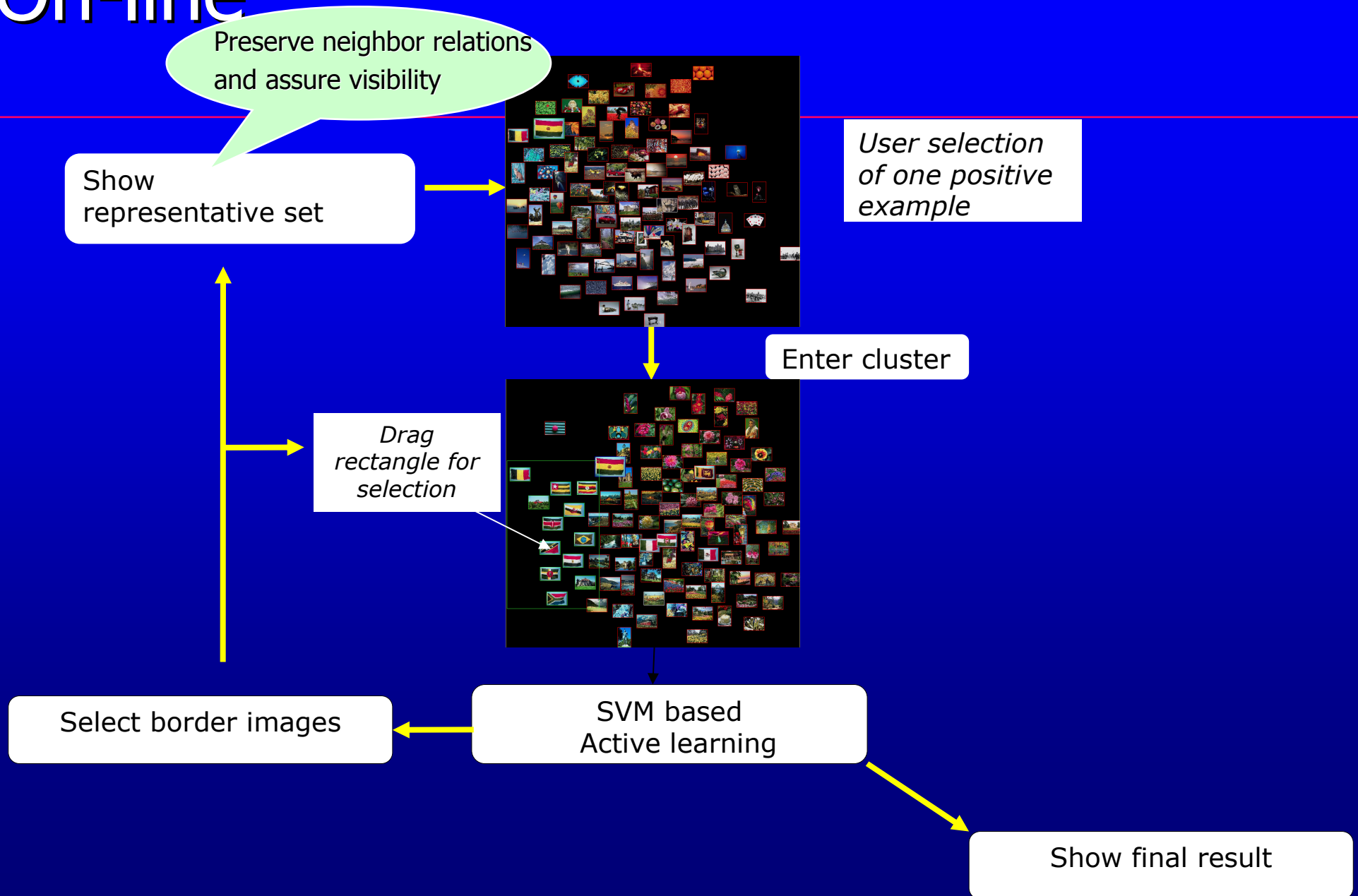
Preserve neighbor relations and assure visibility

Show representative set

User selection of one positive example

Enter cluster

Drag rectangle for selection

Select border images

SVM based Active learning

Show final result

The Galaxy Browser
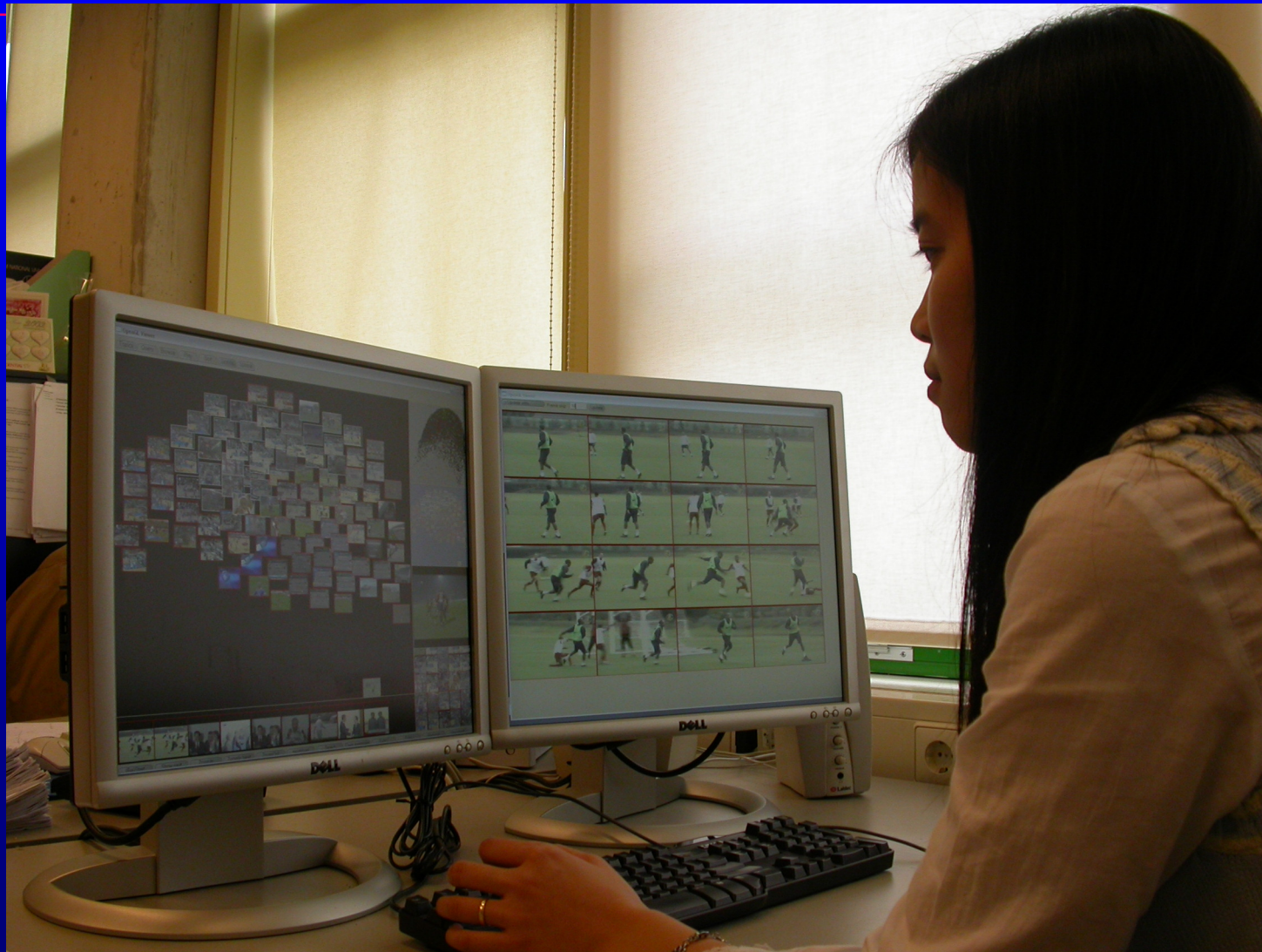
Induced by similarity

User simulation studies show its effectiveness

An effective setup
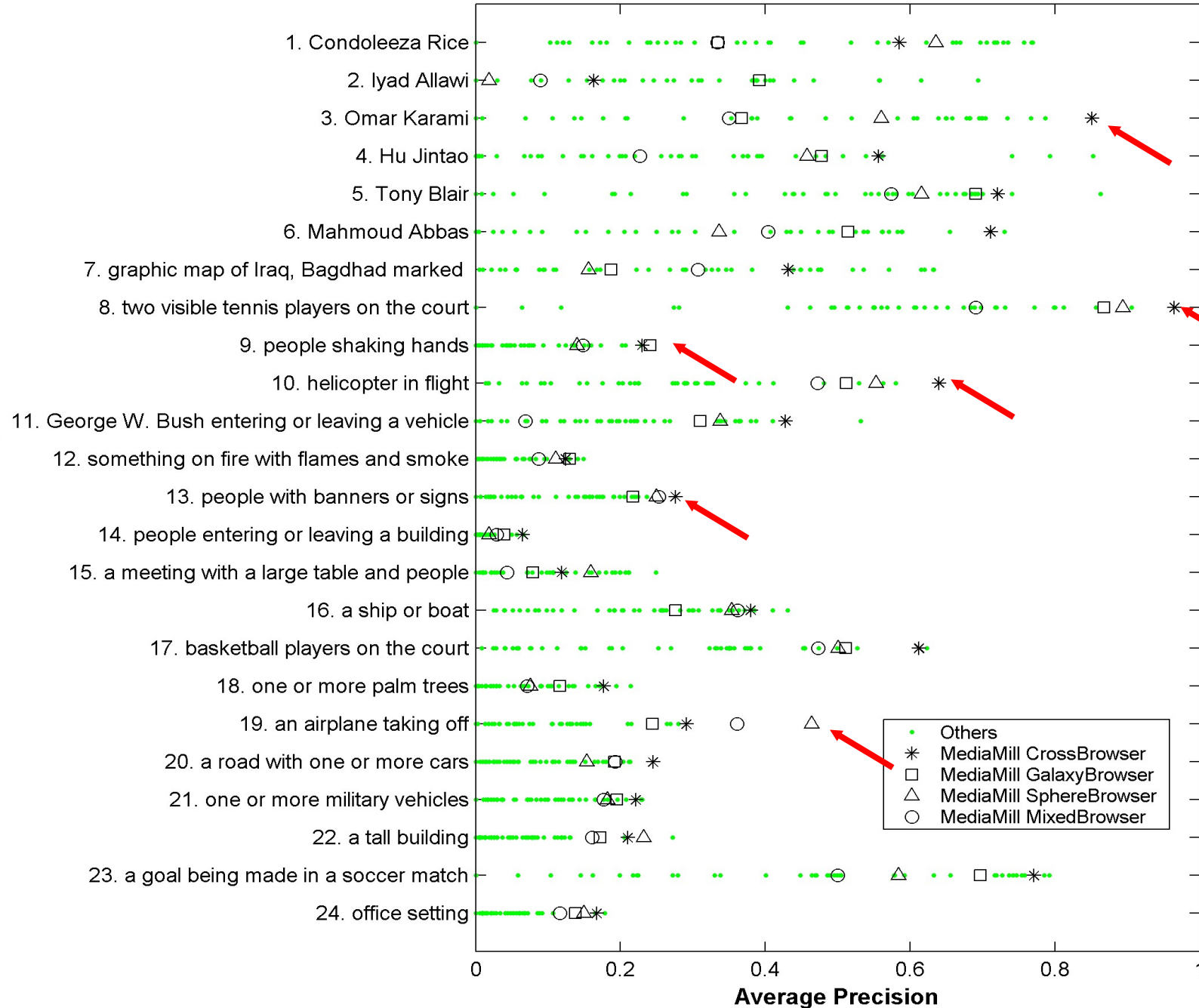
Results

TRECVID 2005 Overall Search Results

Interactive Search Results

Search Topic:
1. Condoleeza Rice
2. Iyad Allawi
3. Omar Karami
4. Hu Jintao
5. Tony Blair
6. Mahmoud Abbas
7. graphic map of Iraq, Bagdhad marked
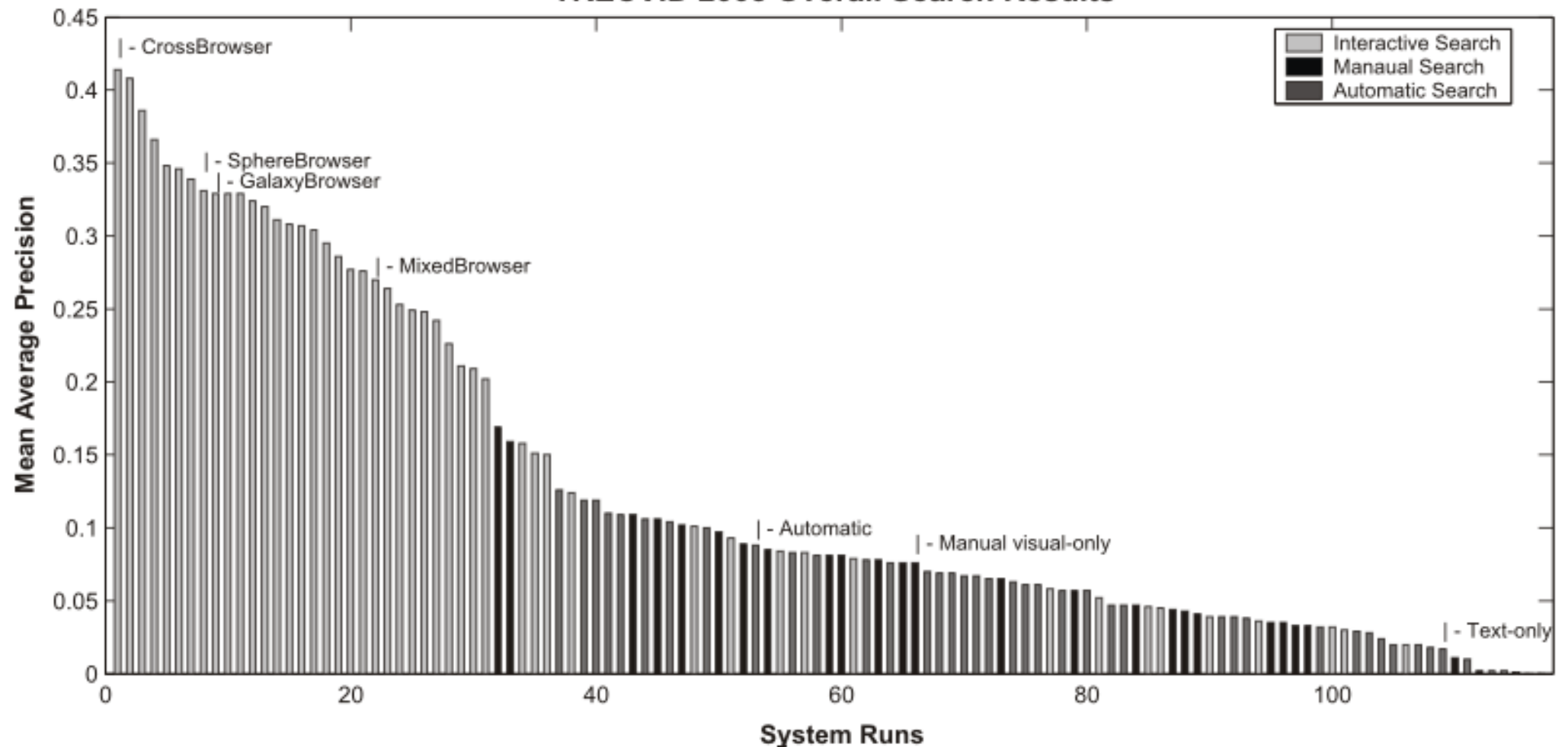8. two visible tennis players on the court
9. people shaking hands
10. helicopter in flight
11. George W. Bush entering or leaving a vehicle
12. something on fire with flames and smoke
13. people with banners or signs
14. people entering or leaving a building
15. a meeting with a large table and people
16. a ship or boat
17. basketball players on the court
18. one or more palm trees
19. an airplane taking off
20. a road with one or more cars
21. one or more military vehicles
22. a tall building
23. a goal being made in a soccer match
24. office setting

Legend:
· Others
* MediaMill CrossBrowser
□ MediaMill GalaxyBrowser
△ MediaMill SphereBrowser
○ MediaMill MixedBrowser

Average Precision (0, 0.2, 0.4, 0.6, 0.8, 1)

Conclusions

- A large multimodal lexicon
  - Is a great aid in all search tasks

- Visualization
  - Helps significantly when there is no 1-1 mapping from query to concept

- Future work
  - Employ semantic concept structure
  - Even more advanced visualizations