

The Problem: Search in video

● What is search?: enter query... → browse results

The screenshot shows the MediaMill Semantic Video Search Engine interface. It features a menu bar with options: F1: Topics, F2: Query, F3: Browse, F4: Sort, Stills, and Help. The main interface is divided into two sections:

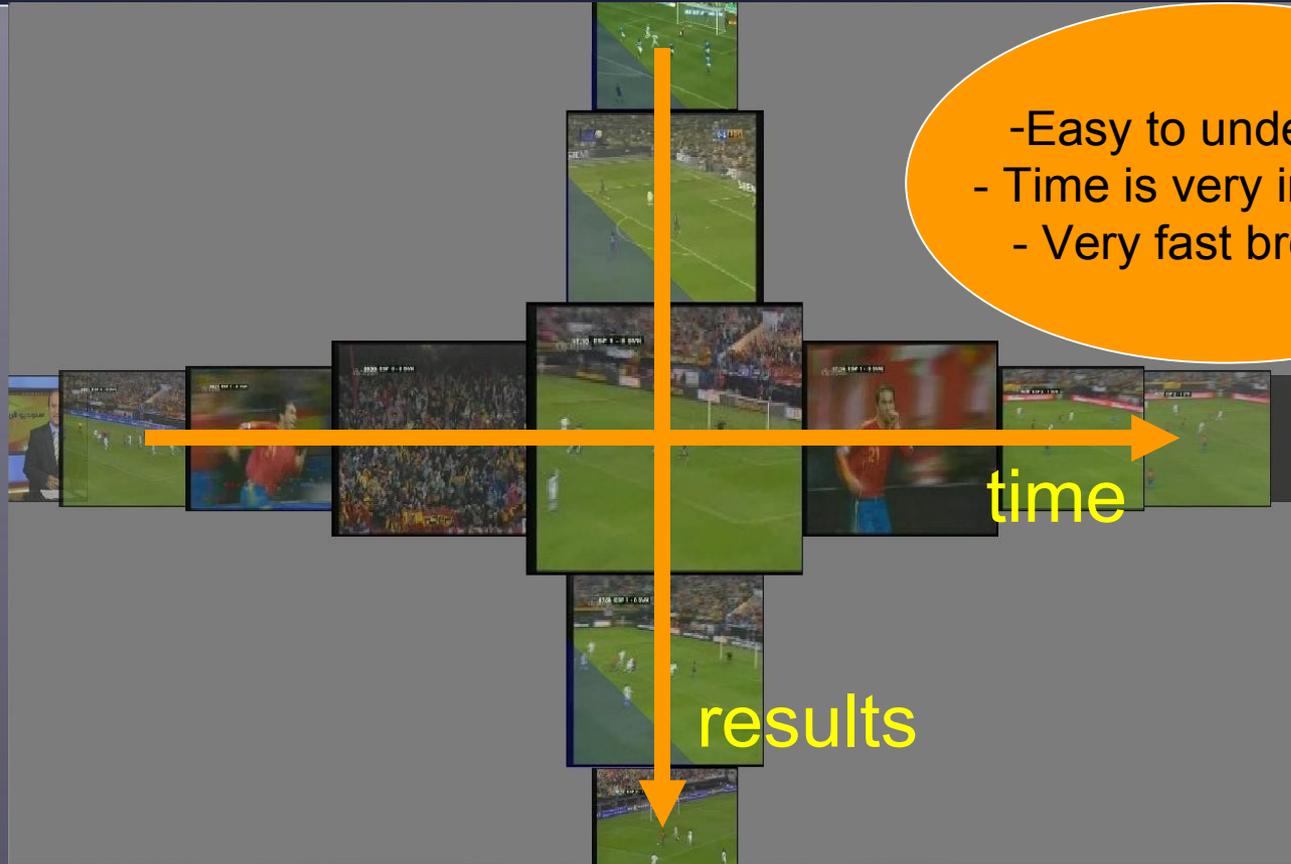
- 1. Define textual query:** Includes a text input field with the placeholder "Find shots of one or more...", a "Query by keyword" label, and a "Preview" window showing "Query by region selection". There are also buttons for "Automated Concept suggestion" and "Query suggestion".
- 2. Add relevant semantic concepts:** Includes a "concept filter:" input field and a list of concepts with associated video thumbnails and progress bars. One concept is "soccer" with a 71% progress bar. Another is "017_soccer" with a 72% progress bar. A "Query component 'mixer'" is also visible.

Annotations on the screenshot include:

- A blue arrow pointing from the "Query by keyword" field to the "Repeat" circle.
- A purple circle containing the text "Repeat Leads to: RANKING".
- A blue arrow pointing from the "Repeat" circle to the "Query component 'mixer'" area.
- Text labels: "Query by example" (pointing to a video thumbnail), "Query by keyword" (pointing to the search input), "Query by region selection" (pointing to the preview window), "Query component 'mixer'" (pointing to the concept list), and "language filter" (pointing to the bottom controls).



Mixing dimensions: CrossBrowser



- Combination of query results and time

Mixing dimensions: RotorBrowser



- Allows combinations based on query on demand

Mixing dimensions

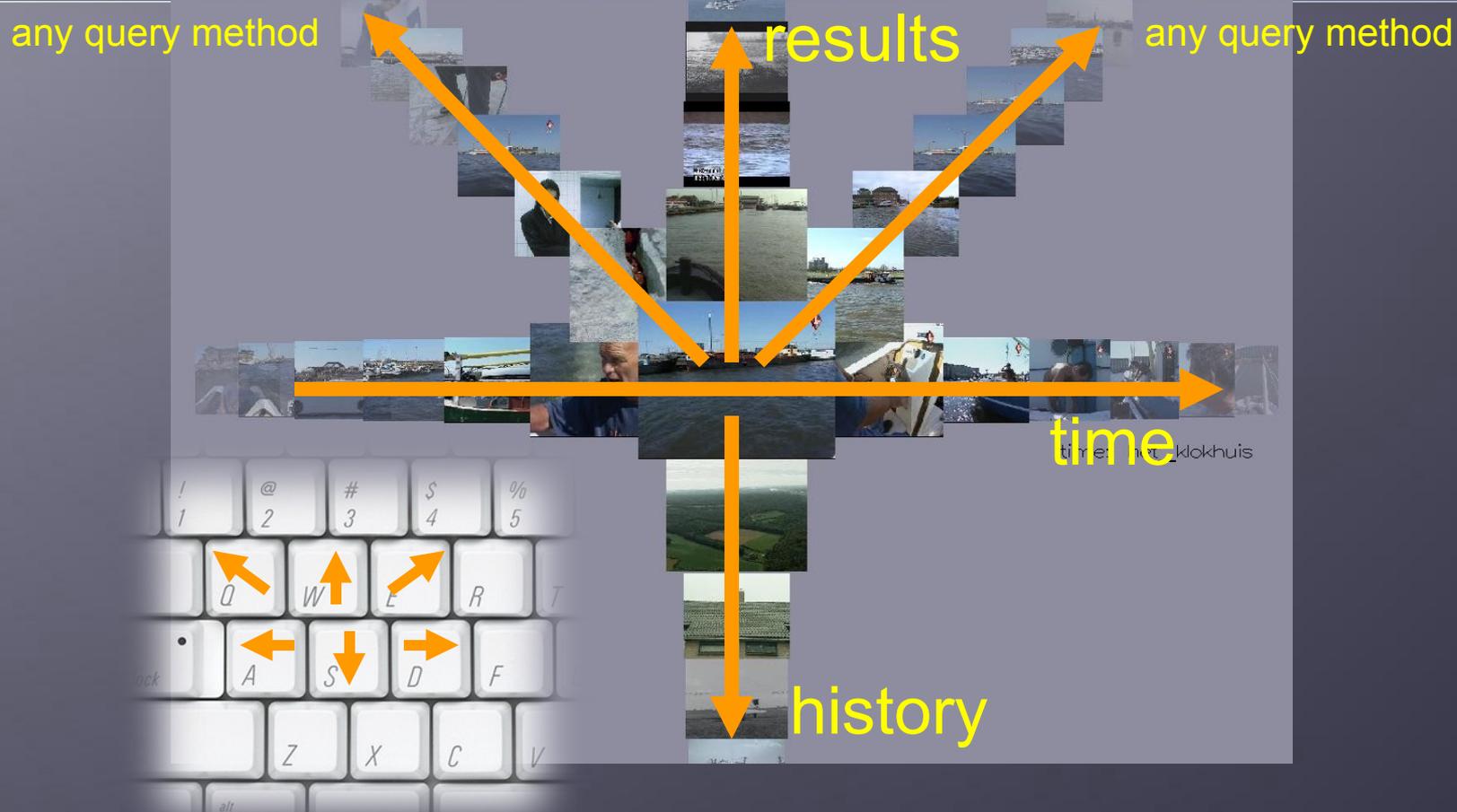


- Easy to understand
- Time is very important
- Very fast browsing
- Doesn't require the user to think of options.
- Allows dataset exploration
- Browsing 'outside query' very important for some types of query
- Limits visits to a "query screen"

Typically a user wants to explore a dataset fast and easy, without difficult query screens.

A hybrid between both browsers is required

The ForkBrowser



● Combination of fixed set of query methods

The ForkBrowser

Combines

- Fast browsing through results
- Assignable dimensions, e.g.
 - Visual similarity
 - semantic similarity
- User doesn't have to revisit query screen
- Animations on demand



The ForkBrowser

The screenshot displays the ForkBrowser interface with the following components:

- Navigation Bar:** F1: Topics, F2: Query, F3: Browse, F4: Sort, Threads, Grid, Help. A red progress bar on the right shows 1414 items.
- 1. Define textual query:** Query: Find shots of a train in motion. Includes radio buttons for 'concept suggestions', 'textual search' (checked), and 'use Dutch engine'. A 'Query' button is present.
- Examples:** A horizontal strip of video thumbnails showing train scenes. A progress bar below indicates 100% completion.
- 2. Add relevant semantic concepts:** A 'concept filter' input field and a list of checkboxes: Concept tree browser, TRECVID_Annotations, MediaMill_Annotations, LSCOM_Annotations, and Color.
- 3. Filter by program:** A list of program titles with counts, including 'andere tijden (3)', 'babylon (1)', 'bij lofbid (1)', 'de adriekant van het oo', 'de kleine wereld (1)', 'duitsch direkt (1)', 'documentaires (1)', 'dokwerk (7)', 'dordrecht, holland's eeltes', 'eerste druk met midas (1)', 'feest (1)', 'feesten met een verhaal', 'fernhout, de filmer (1)', 'geo topics (1)', 'gare d'italia (1)', 'hollige huisjes (1)', 'herman van der horst win', 'herontdakt (1)', 'het kokhus (16)', and 'het laatste zaad (1)'. 'all', 'none', and 'invert' buttons are at the bottom.
- 4. Construct the initial query:** A horizontal strip of video thumbnails with a progress bar at 100%.
- Buttons:** 'Show videos', 'Clear', 'Set supplemental', and 'Search' are located at the bottom of the interface.

Experiments

● TRECVID 2007 Interactive Search

● We compare:

- Run with CrossBrowser (UVA_MM_1)
- Run with ForkBrowser (UVA_MM_2)
- Evaluation metrics try to minimize effect of comparing expert users

● Set up:

- Seed:
 - Automatic search results
 - Query by concept, keyword and example
- Extra 'tines' in ForkBrowser:
 - Weibull and Gabor visual similarity features

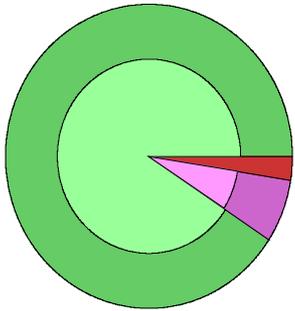
● What do we want to know?

- Is browsing using multiple dimensions useful?
- Does a fixed layout lead to faster browsing and better results?

- Is browsing multiple dimensions useful?
 - Evaluate effectiveness of having multiple dimensions

Query method usage per topic

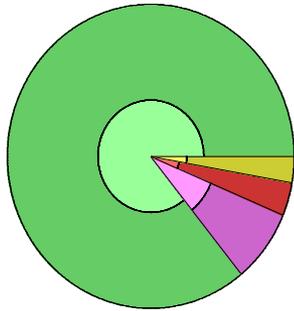
0205



train in motion

Helped by:
animation
time

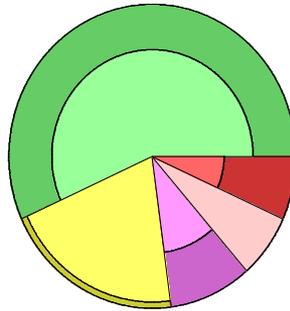
0212



boat moving past

Helped by:
animation
time

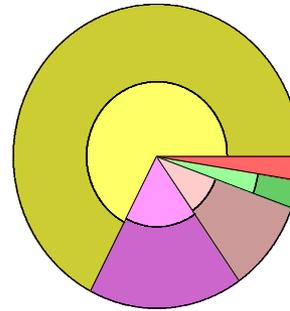
0209



3+ people at table

Helped by:
concept
time
visual similarity

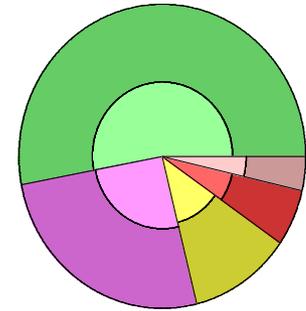
0213



woman talking

Helped by:
visual similarity
concept

0214

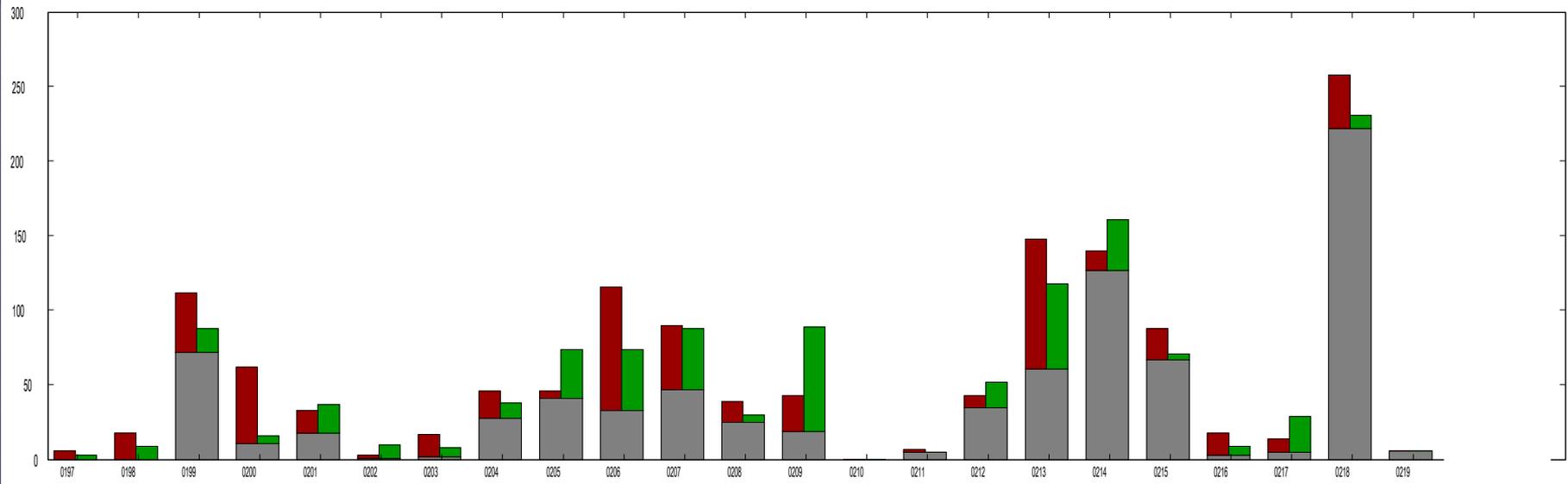


large crowd

Helped by:
concept

Different topics
have different search
strategies

Unique results per browser

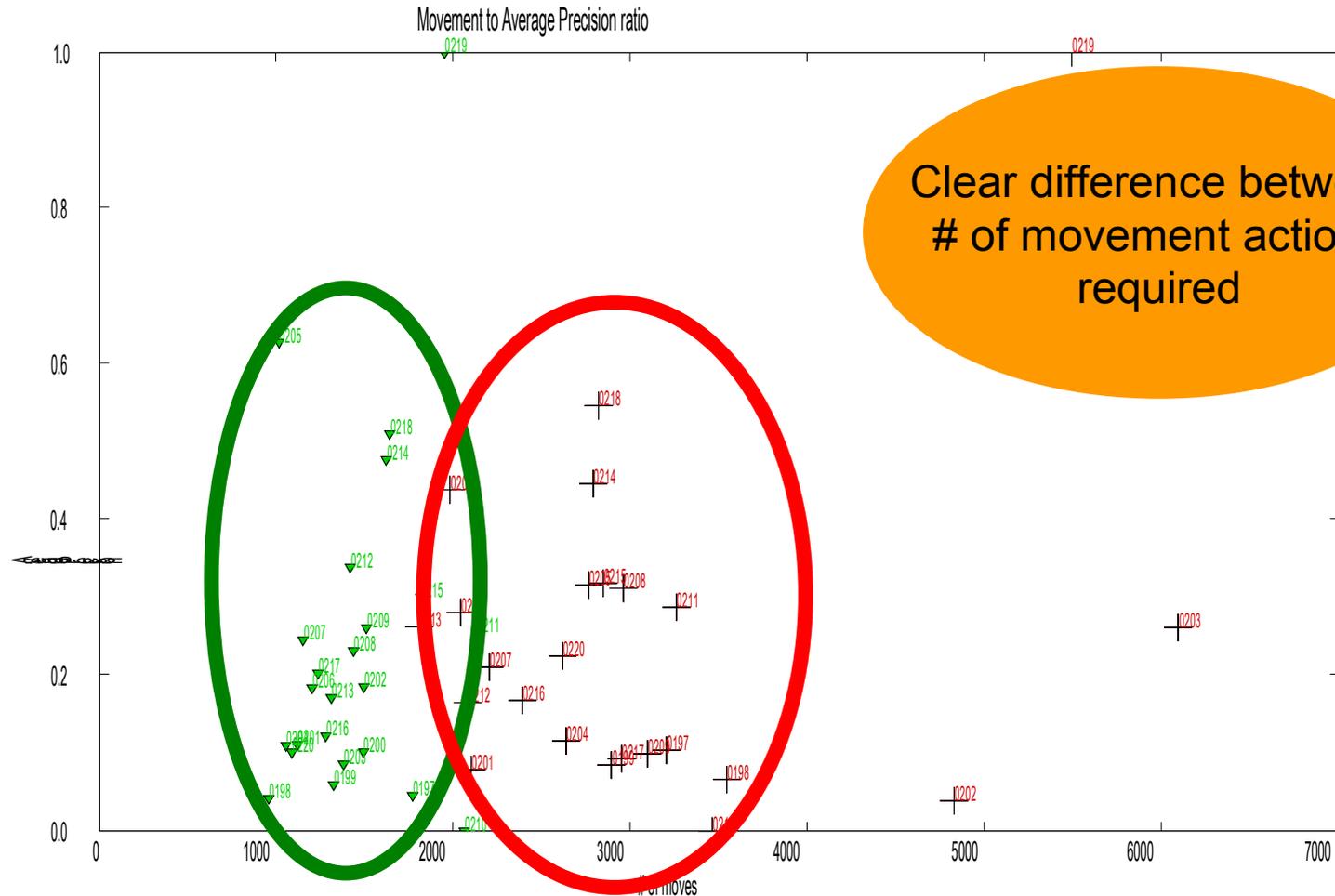


Graph shows the number of (correct) shots retrieved

- Grey:** shots found by both browser
- Red:** shots found only by the CrossBrowser
- Green:** shots found only by the ForkBrowser

Both browsers find different results

Movement vs Average Precision



Conclusions

● Evaluation:

- Different combinations of query dimensions are beneficial for individual topics
- ForkBrowser requires less interaction steps from the user for the same average precision
- Both browsers find different unique results

Any questions?

