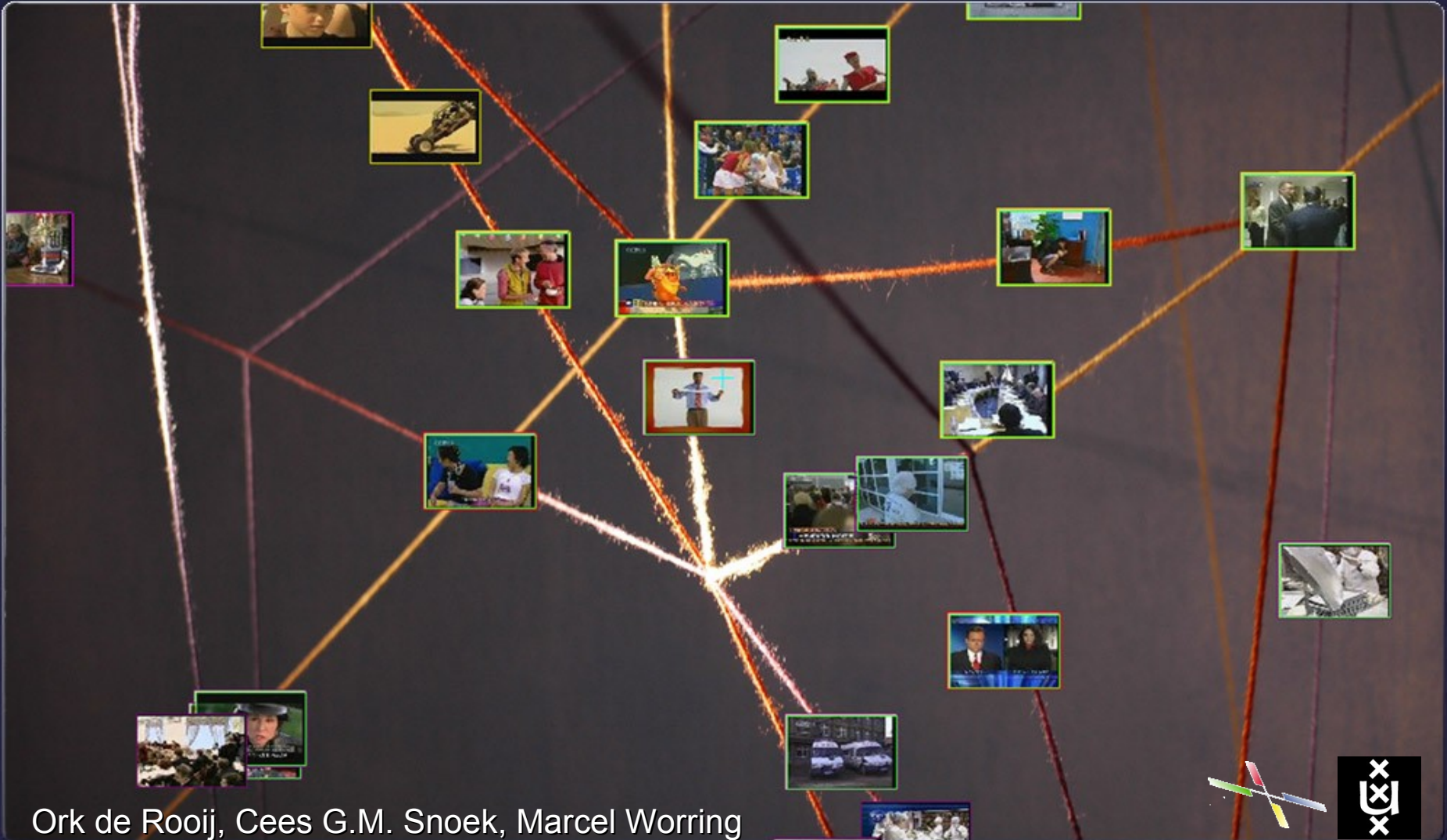


Consuming videos with the ForkBrowser



Ork de Rooij, Cees G.M. Snoek, Marcel Worring



The Problem: Search in video

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

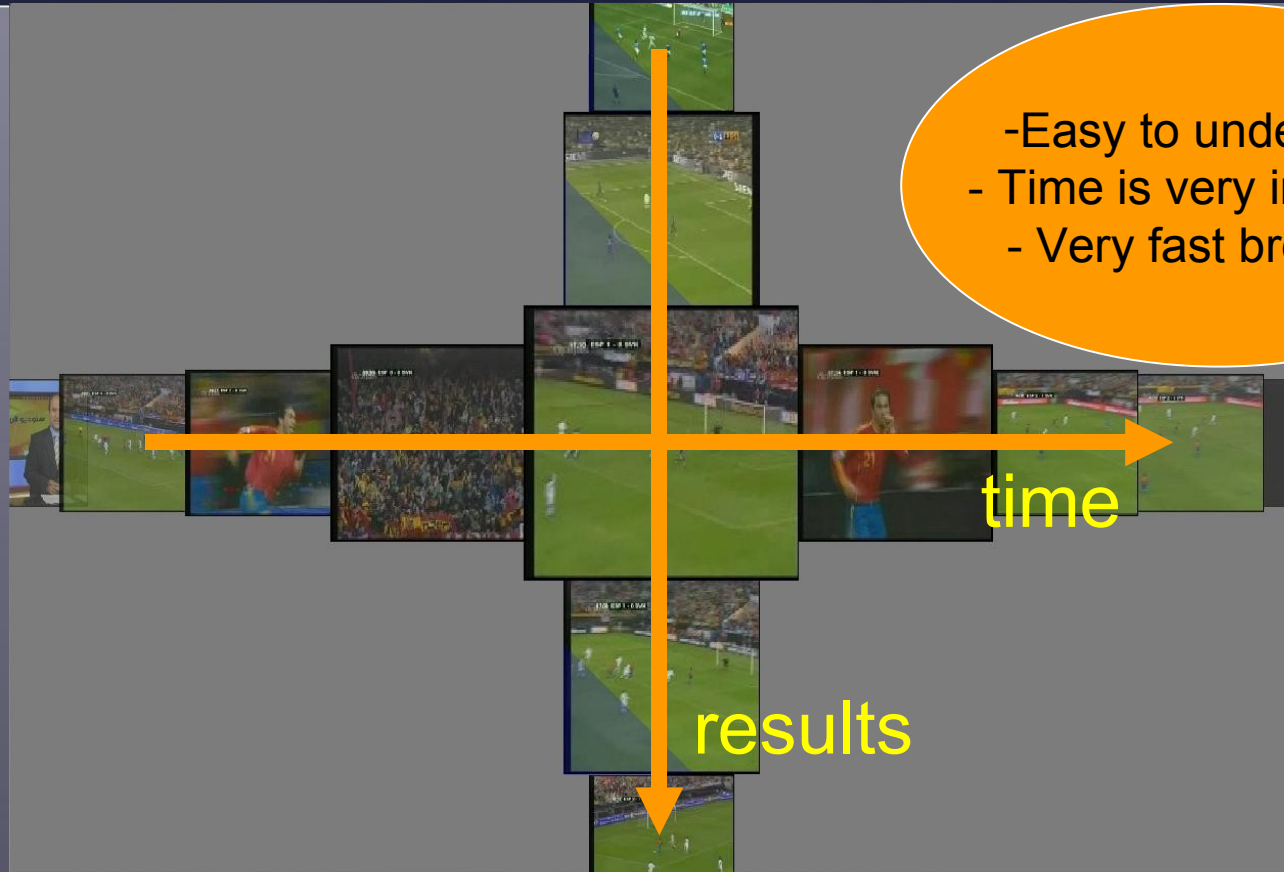
The screenshot shows the MediaMill Semantic Video Search Engine interface. The window title is "MediaMill Semantic Video Search Engine". The interface is divided into several sections:

- 1. Define textual query:** A search bar with the text "Find shots of one or more..." and a "Query by keyword" label. Below it are "Examples" and "Automated Concept suggestion" buttons.
- 2. Add relevant semantic concepts:** A section with a "concept filter:" input field and a "Repeat" button.
- Query by example:** A section showing video thumbnails with a "Query by example" label.
- Query by region selection:** A section with a "Preview" window and "Visual" and "Region" buttons.
- Language filter:** A section with a "language filter" label and a "Show videos" button.
- Search and Ranking:** A section with a "Search" button and a "Leads to: RANKING" label.

Annotations include a large blue "Repeat" arrow pointing to the "Repeat" button, a purple oval containing the text "Leads to: RANKING", and a white oval containing the text "Query component 'mixer'".

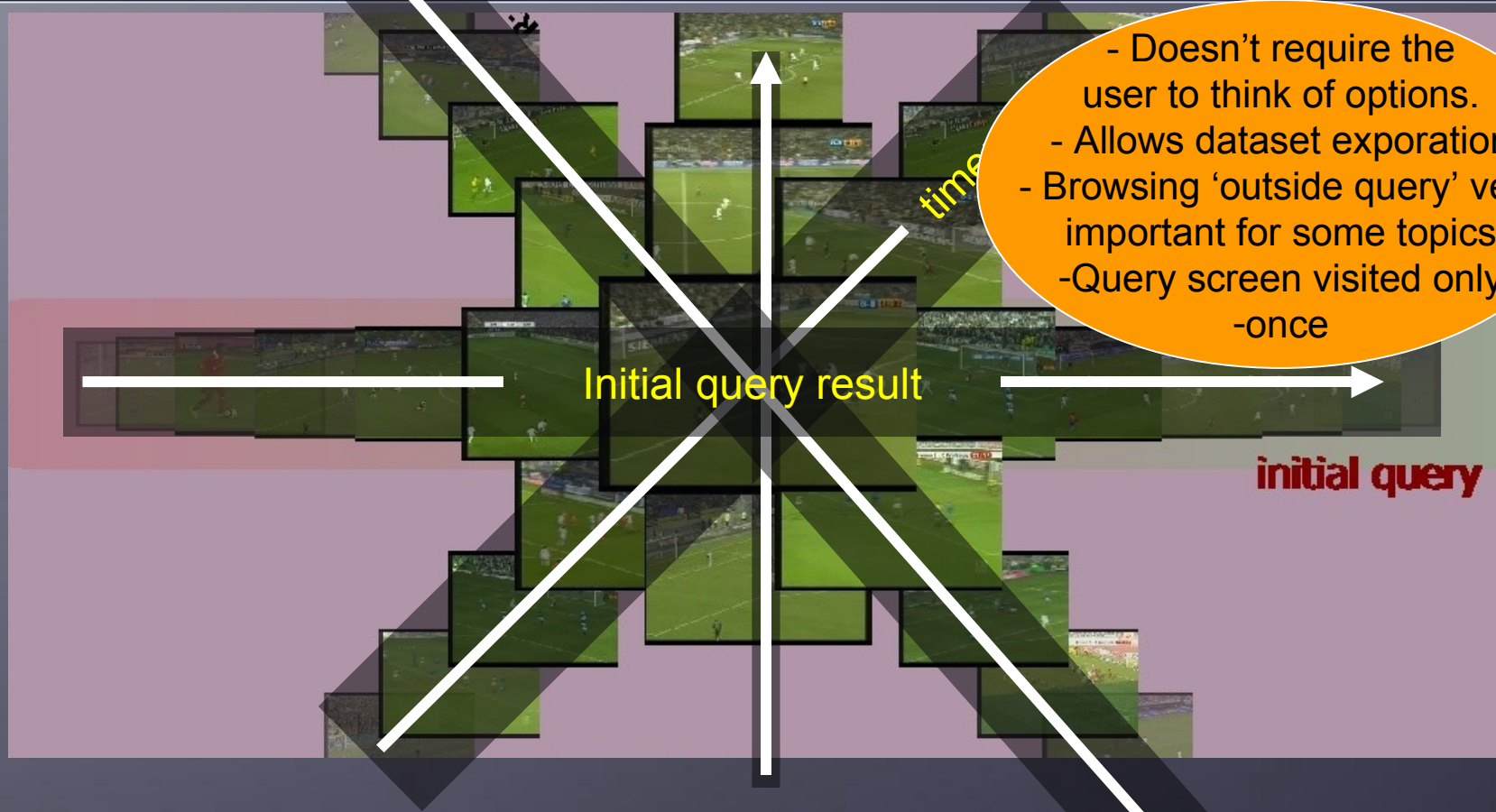


Mixing dimensions: CrossBrowser



- Combination of query results and time

Mixing dimensions: RotorBrowser



- Allows combinations based on query on demand

Mixing dimensions

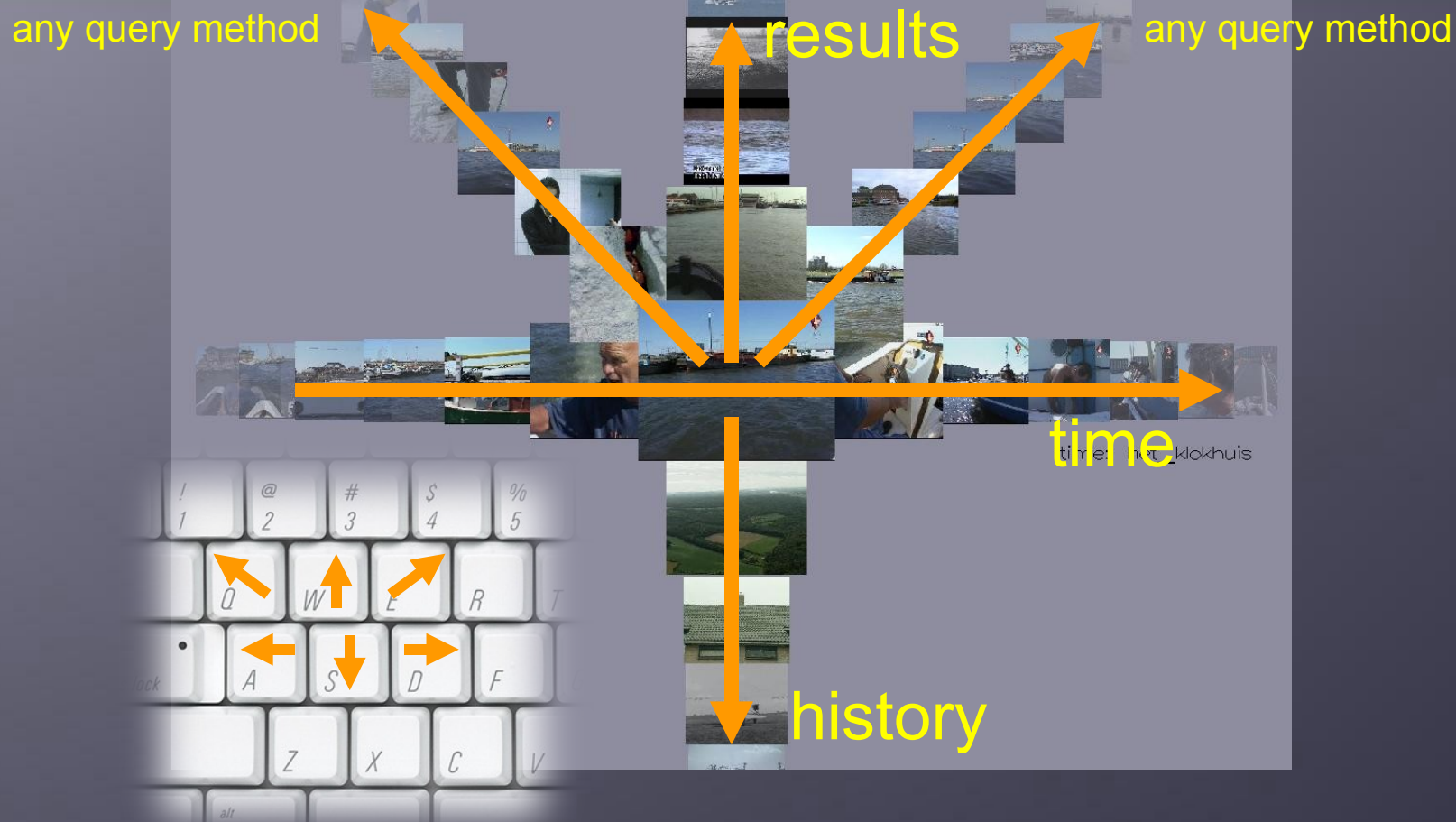


- Easy to understand
- Time is very important
- Very fast browsing
- Doesn't require the user to think of options.
- Allows dataset exploration
- Browsing 'outside query' very important for some types of query
- Limits visits to a "query screen"

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

A hybrid between both browsers is required

The ForkBrowser

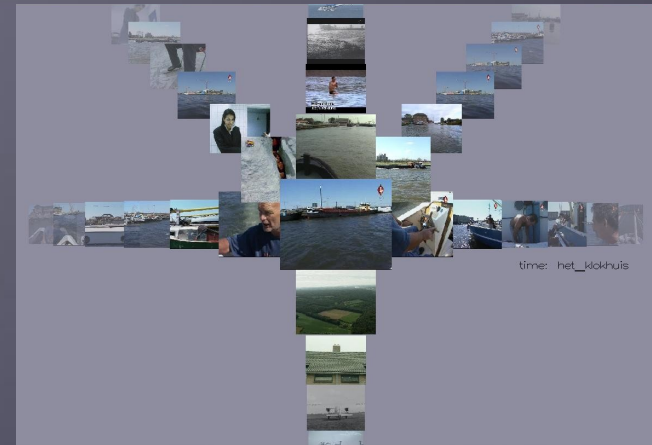


● Combination of fixed set of query methods

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 tree view under 'Concept tree browser' with items: TRECVID_Annotations, MediaMIL_Annotations, LSCOM_Annotations, and Color.
- 3. Filter by program:** A list of program titles such as 'andere tijden (3)', 'babylon (1)', 'bij lofbid (1)', 'de adriekant van het oer', '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)', 'halloze huisjes (1)', 'herman van der horst win', 'herontdakt (1)', 'het kokhus (16)', and 'het laatste zood (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:

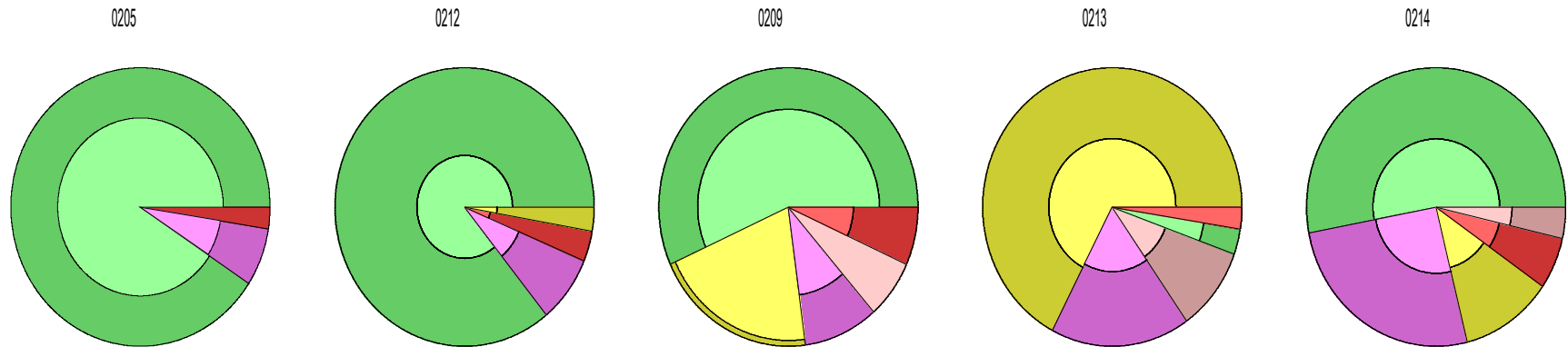
- Seed:
 - Automatic search results
 - Query by concept, keyword and example
- Extra 'tines' in ForkBrowser:
 - Weibull and Gabor visual similarity features

● 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



train in motion

boat moving past

3+ people at table

woman talking

large crowd

Helped by:
animation
time

Helped by:
animation
time

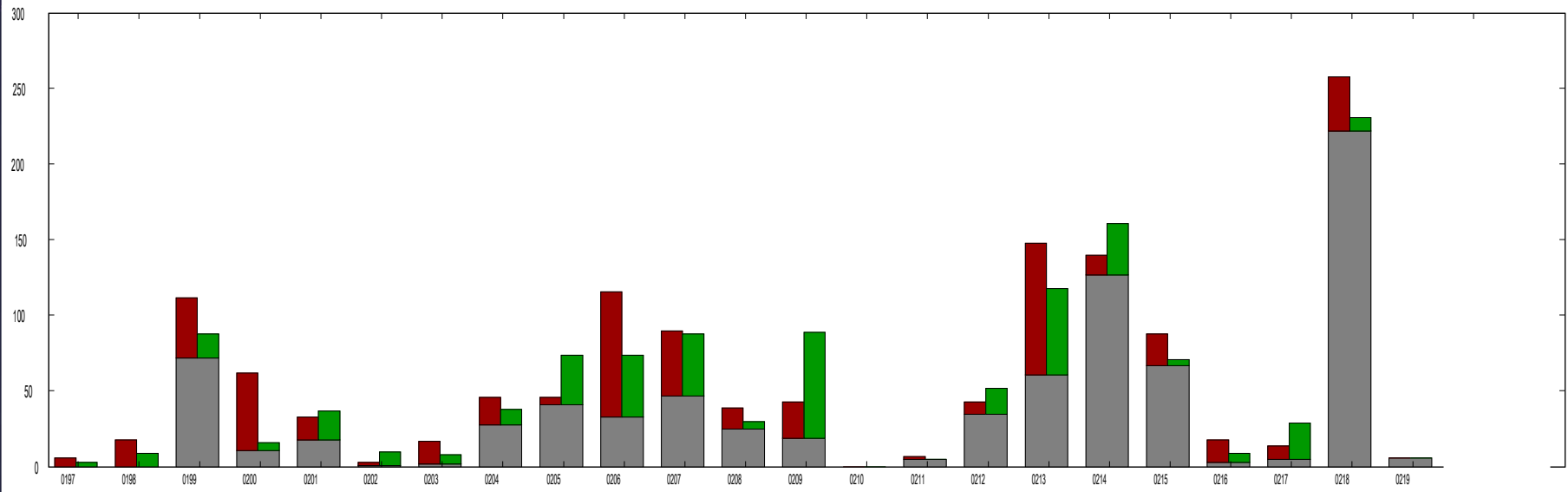
Helped by:
concept
time
visual similarity

Helped by:
visual similarity
concept

Helped by:
concept

Different topics
have different search
strategies

Unique results per browser

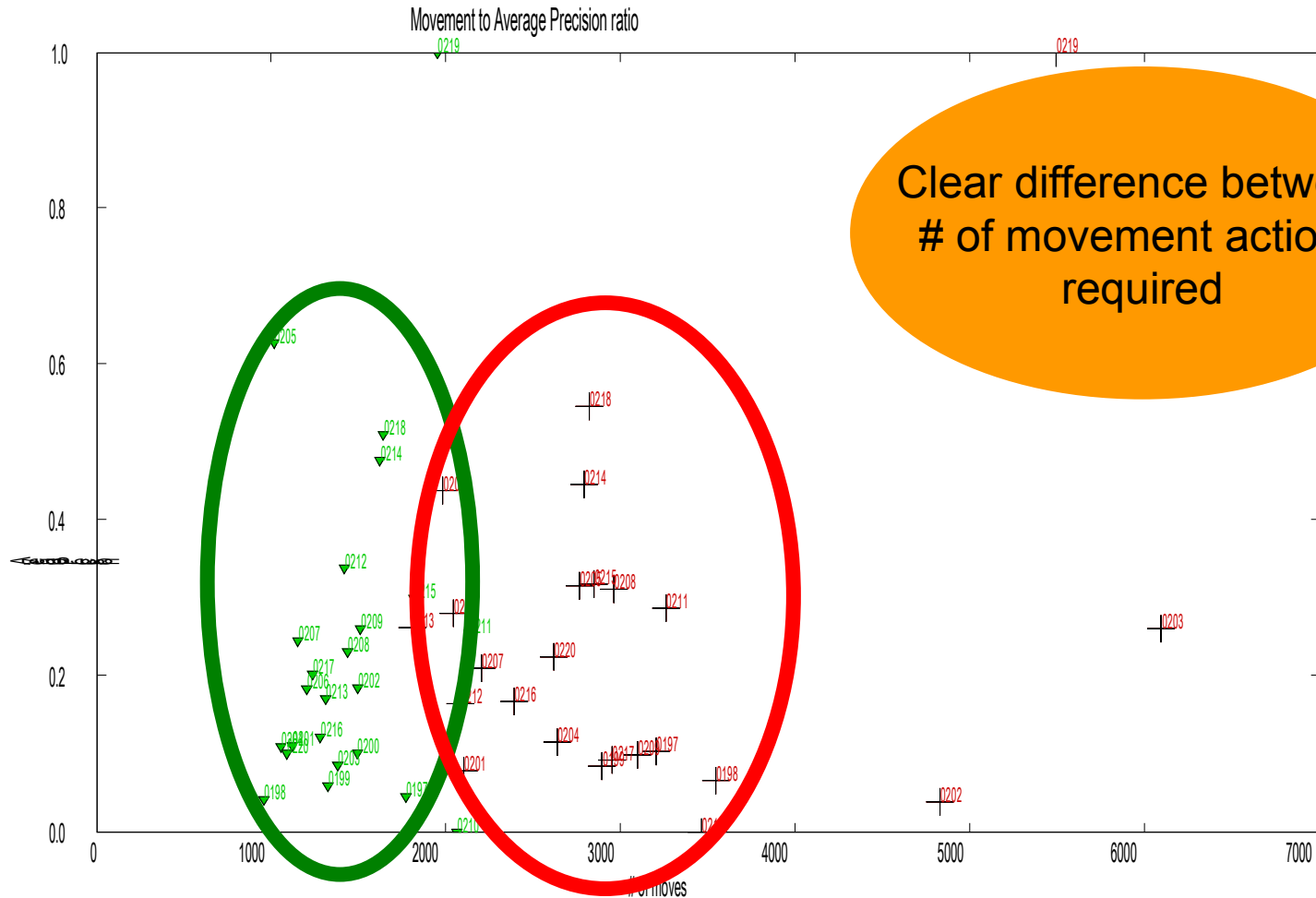


Graph shows the number of (correct) shots retrieved

- 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



Clear difference between
of movement actions
required

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?

