

CDVP & TRECVID-2003

Interactive Search Task Experiments

Paul Browne, Georgina Gaughan, **Cathal Gurrin**,
Gareth J.F. Jones, Hyowon Lee, Sean Marlow,
Kieran Mc Donald, Noel Murphy, Noel E.
O'Connor, Alan F. Smeaton, Jiamin Ye

Centre for Digital Video Processing
Dublin City University, Glasnevin, Dublin 9, Ireland

Contents

- Introduction
 - Físchlár Systems
- Interactive Search Experiment
 - System & Experiment Design
 - System Demonstration
 - Submitted Runs
- Findings
 - Comparing Systems Performance
 - User Observations
- Conclusions

Físchlár Demonstrator System

- A Digital Video Management System
 - Web-based, supports browsing and search
 - Many different versions of the system
 - Underlying XML Architecture
 - XSL supporting display on multiple devices
- TREC2003 is our 3rd TRECVID Search Task
 - **2003** : explored benefits of incorporating image and feedback into a text search process
 - **2002** : explored benefits of incorporating features
 - **2001** : examined different keyframe browsers

Interactive Search Experiment

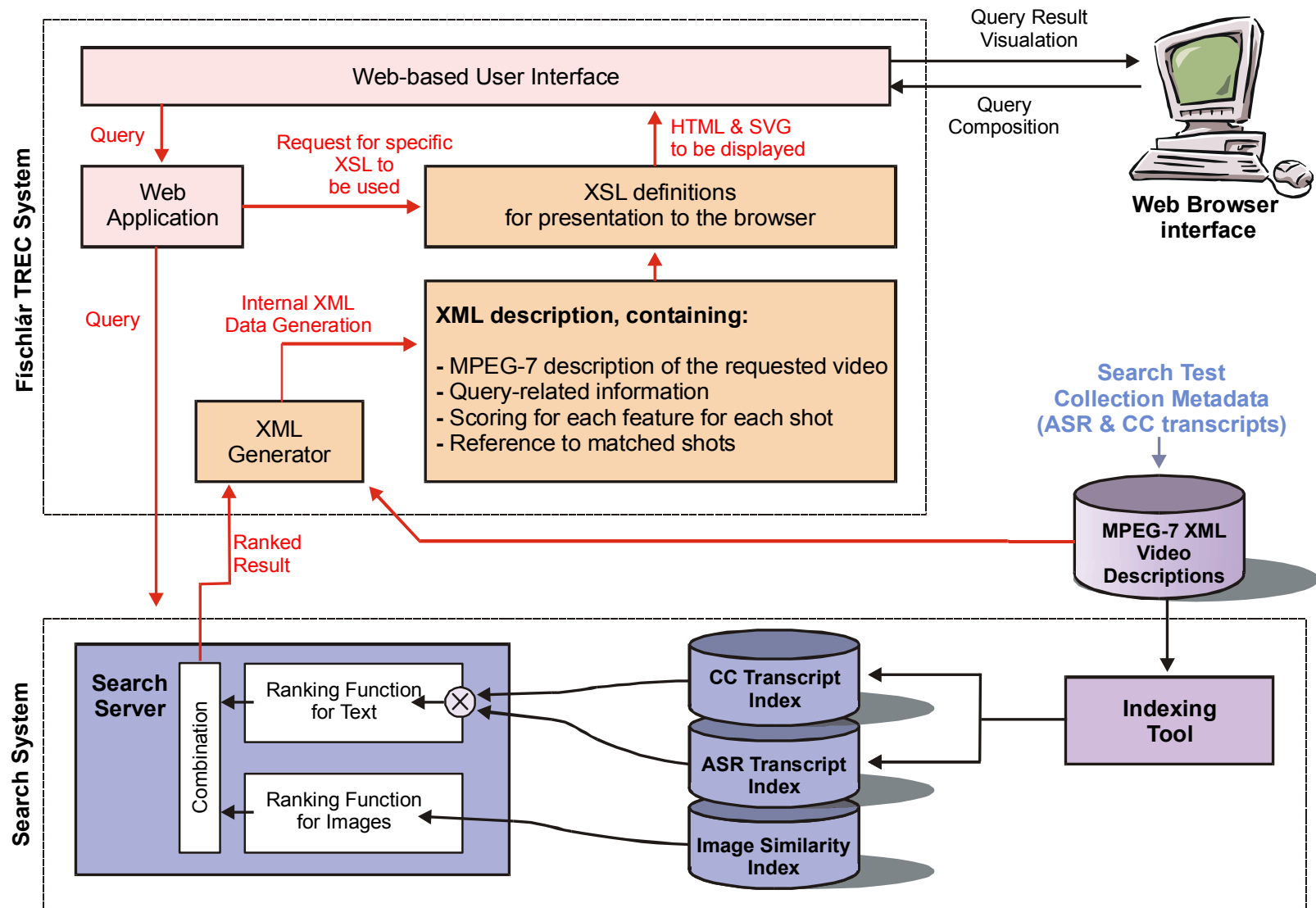
- Testing if a text/image search system incorporating 'more like this' feedback outperforms a text-only system.
- Developed two Físchlár systems:
 - Each highly interactive with a keyframe browser and playback window
 - **(1) Text-only** search and retrieval
 - ASR (LIMSI) & CC Text
 - **(2) Text & Image** search incorporating a **feedback** mechanism
 - ASR & CC Text
 - Keyframe-keyframe similarity (image matching)
 - 'more like this' feedback

Experiment Set-up

- User experiments in a computer lab environment
 - We used the recommended mixing algorithm for searchers / topics
 - Number of Users : 16
 - Typical postgraduate students
 - No prior experience of using the system
 - Topics per User : 12 (6 per system)
 - Minutes per Topic : 7 (last year 4 mins)
 - Each topic evaluated 8 times, 4 times on each system... reduces the effect of user variability
 - Users were trained for 10 mins then allowed two sample topics before experiment
 - Coffee, cookies & headphones were provided



System Architecture

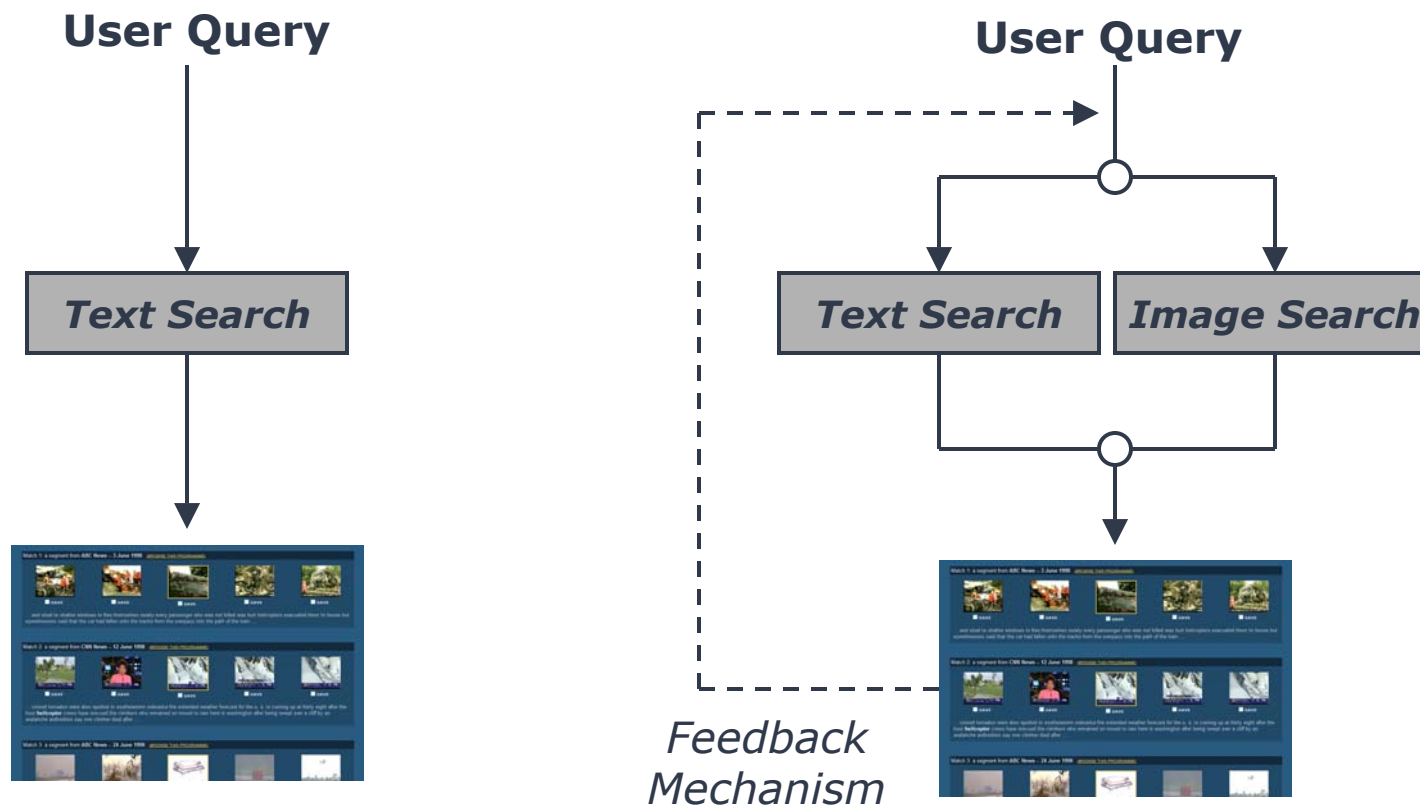


Two search options

- Text Search
 - Using conventional Search Engines (BM25)
 - Two employed, simple combination:
 - ASR Text
 - CC Text
 - Required alignment with the ASR text
- Image Search
 - Keyframe-keyframe or query image-keyframe similarity using:
 - 4 low-level visual features
 - 3 colour-based features and 1 edge-based feature
 - Combined to produce dis-similarity values and were then normalised

User Interaction Differences

- User Interaction is/can be different for both systems:



Format of Results

- Results presented as Groups of Shots
 - Five sequential shots
 - Associated ASR text is also presented



- Each shot contributes to the overall score of the group (0.08, 0.16, 0.5, 0.16, 0.08)
- Top 100 groups of shots ranked and presented in pages of size 20

Feedback Mechanism

Query panel

QUERY
Enter terms. Added text & images will be used together for searching. [clear text] [clear all]

yasser arafat Search

Type in search term(s) and
Click on **Search** button

Search result

Match 1: a segment from **ABC News -- 3 May 1998** (BROWSE THIS PROGRAMME)



Below each thumbnail are buttons: ☐ SAVE and **ADD TO QUERY**. The third thumbnail has a checked ☒ SAVE button.

... the players arrived in london all of them downplaying the chances for a breakthrough here israeli prime minister benjamin netanyahu has called the chances for an agreement the stretch the mediators secretary of state albright talked about unbridgeable gaps that before he left for london this morning palestinian chairman **yasser arafat** was no more optimistic four and a half years after the peace agreement was signed the hope inspired has all but disappeared israel has delayed repeatedly with dralls it promised to make from land palestinians want for their state ...


Query panel

QUERY
Enter terms. Added text & images will be used together for searching. [clear text] [clear all]

yasser arafat Search

TEXT <<----- IMPORTANCE ----->> IMAGE

that before he left for london this morning palestinian chairman yasser arafat was no more optimistic four (REMOVE)



Clicking on **Add to Query** button below a keyframe adds that shot content (text and image) into Query panel: subsequent search will use this shot along with the initial text term used

Demonstration

Text, Image & Feedback System Demonstration



14 shots saved ([view saved shots](#))
 >> [Finish this Task](#)

QUERY

Enter terms. Added text & images will be used together for searching. [\[clear text\]](#) [\[clear all\]](#)

TASK 5:

Find shots from behind the pitcher in a baseball game as he throws a ball that the batter swings at

Have a look at the examples below as hint for your search - where image and video examples are shown, you can click on the **ADD TO QUERY** icon to use that image as a part of your query in your searching.

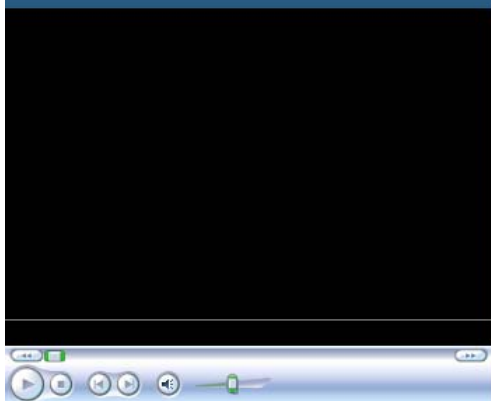
Video example: [\(Play this clip\)](#) ADD TO QUERY

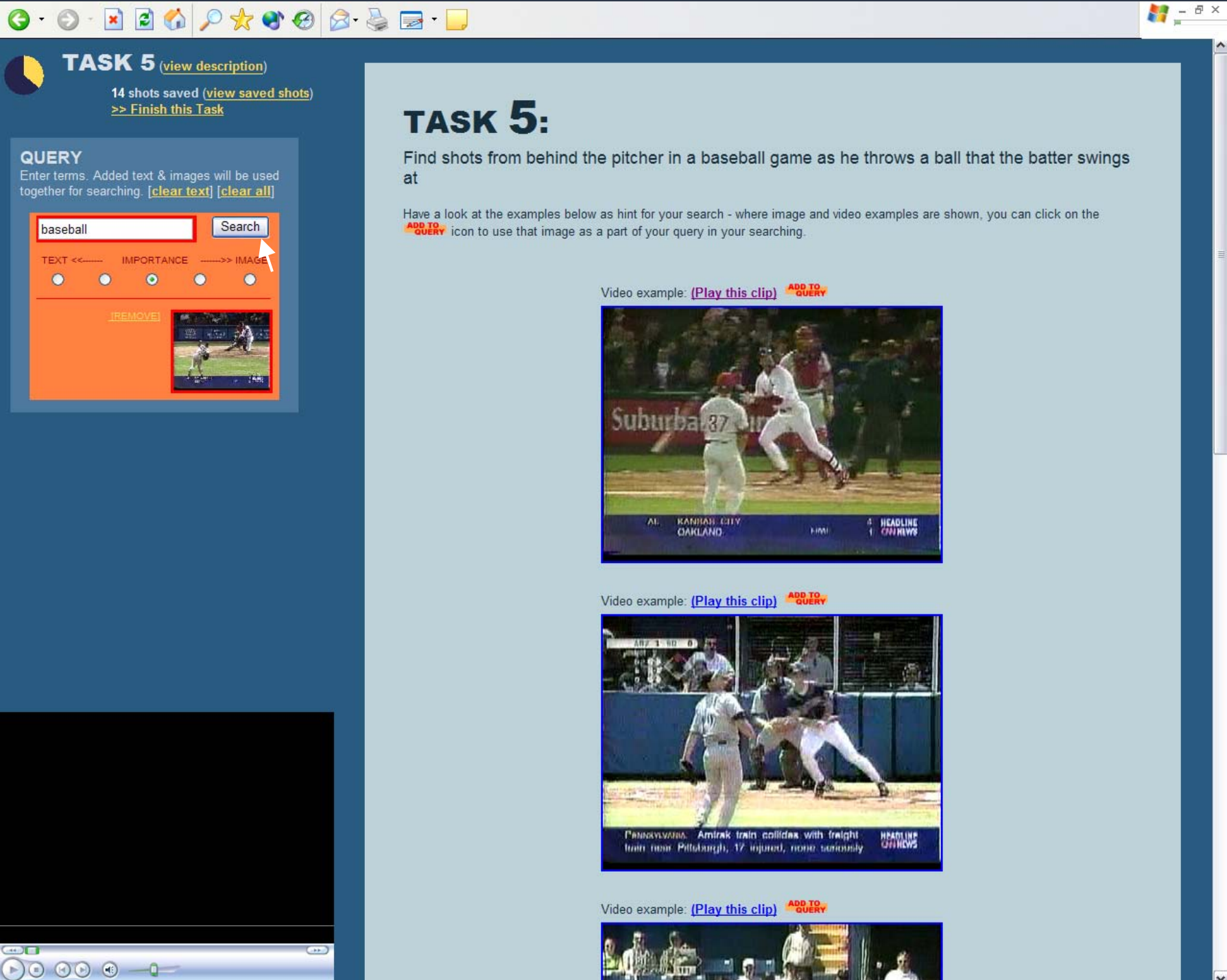


Video example: [\(Play this clip\)](#) **ADD TO QUERY**

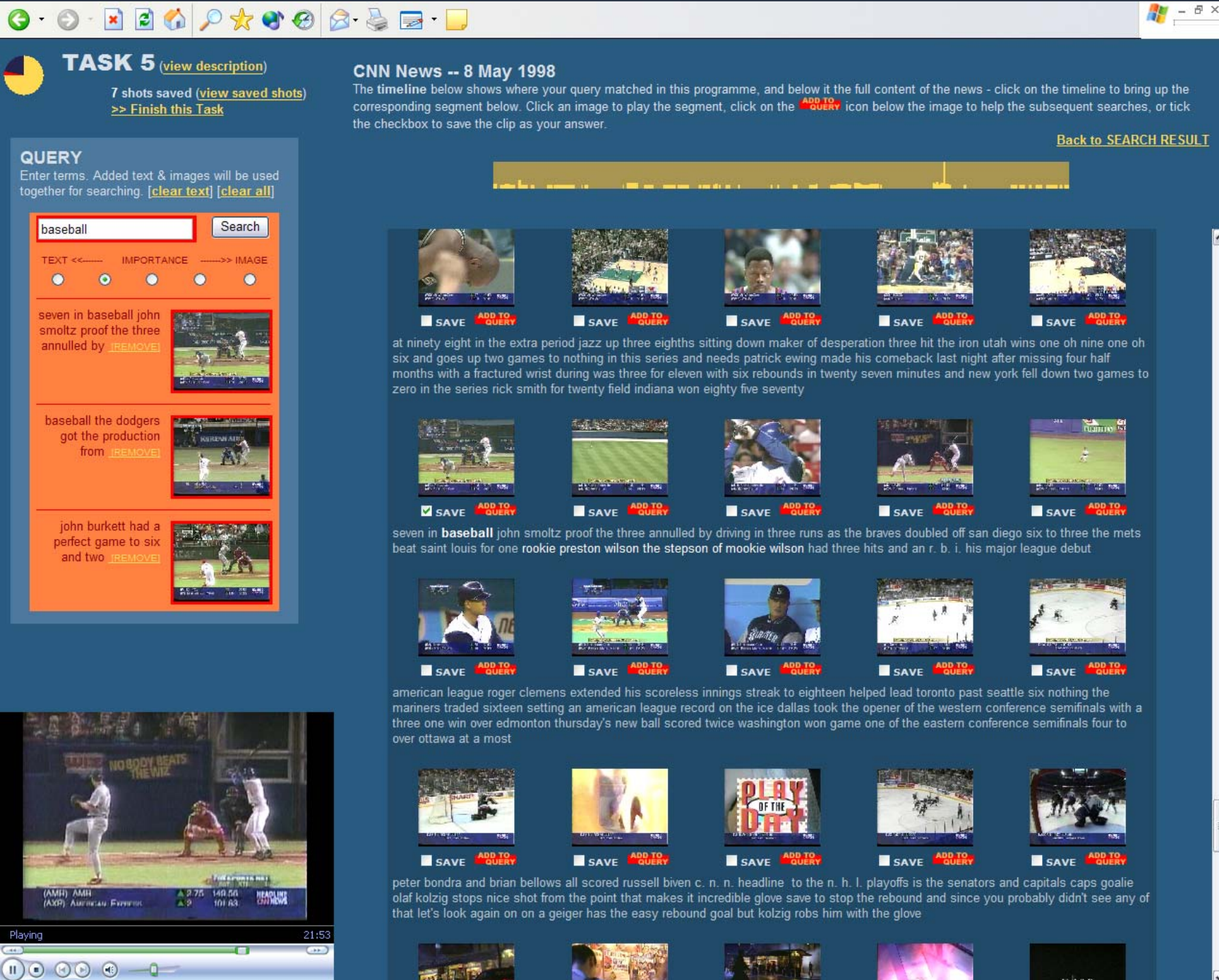


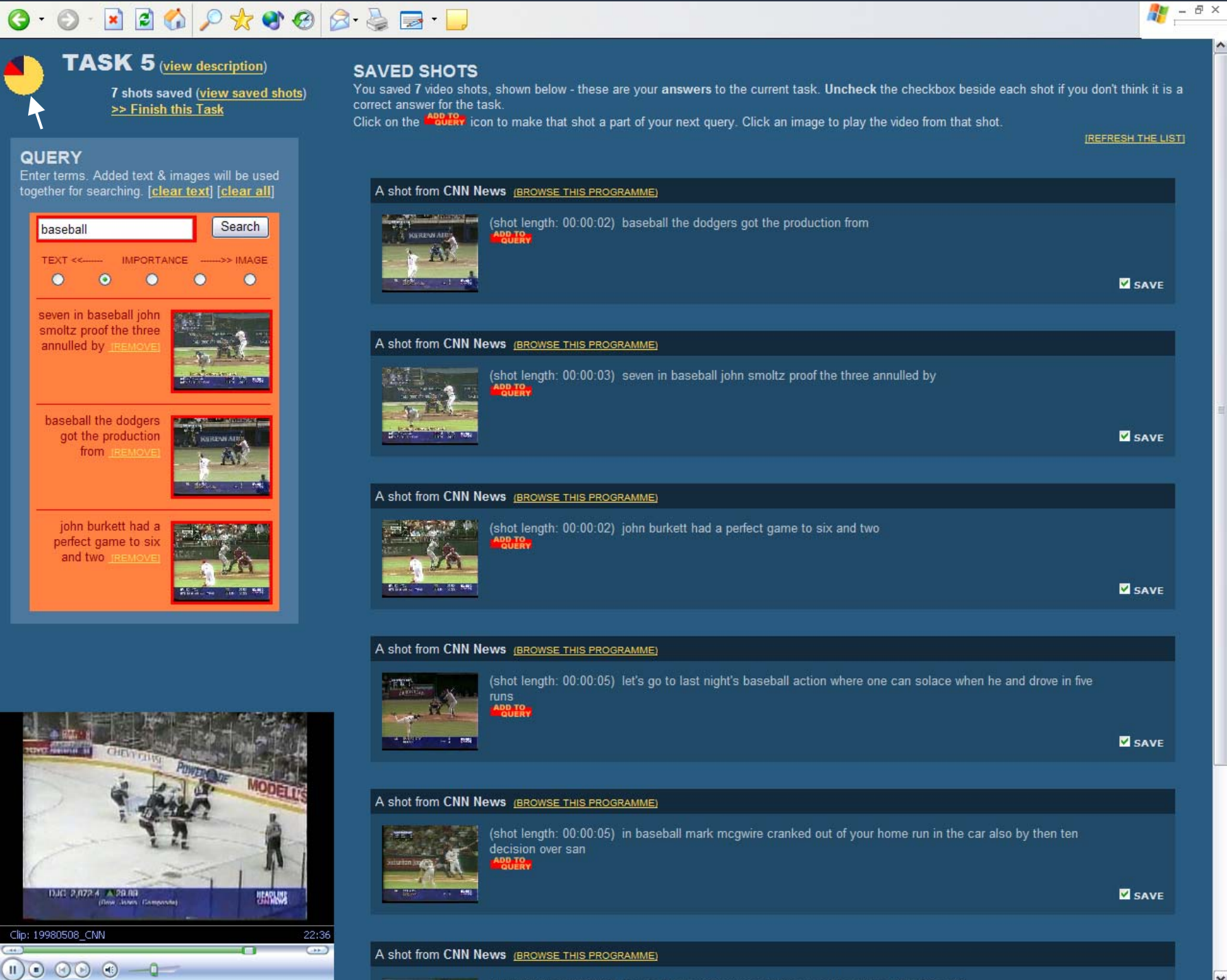
Video example: [\(Play this clip\)](#) **ADD TO QUERY**

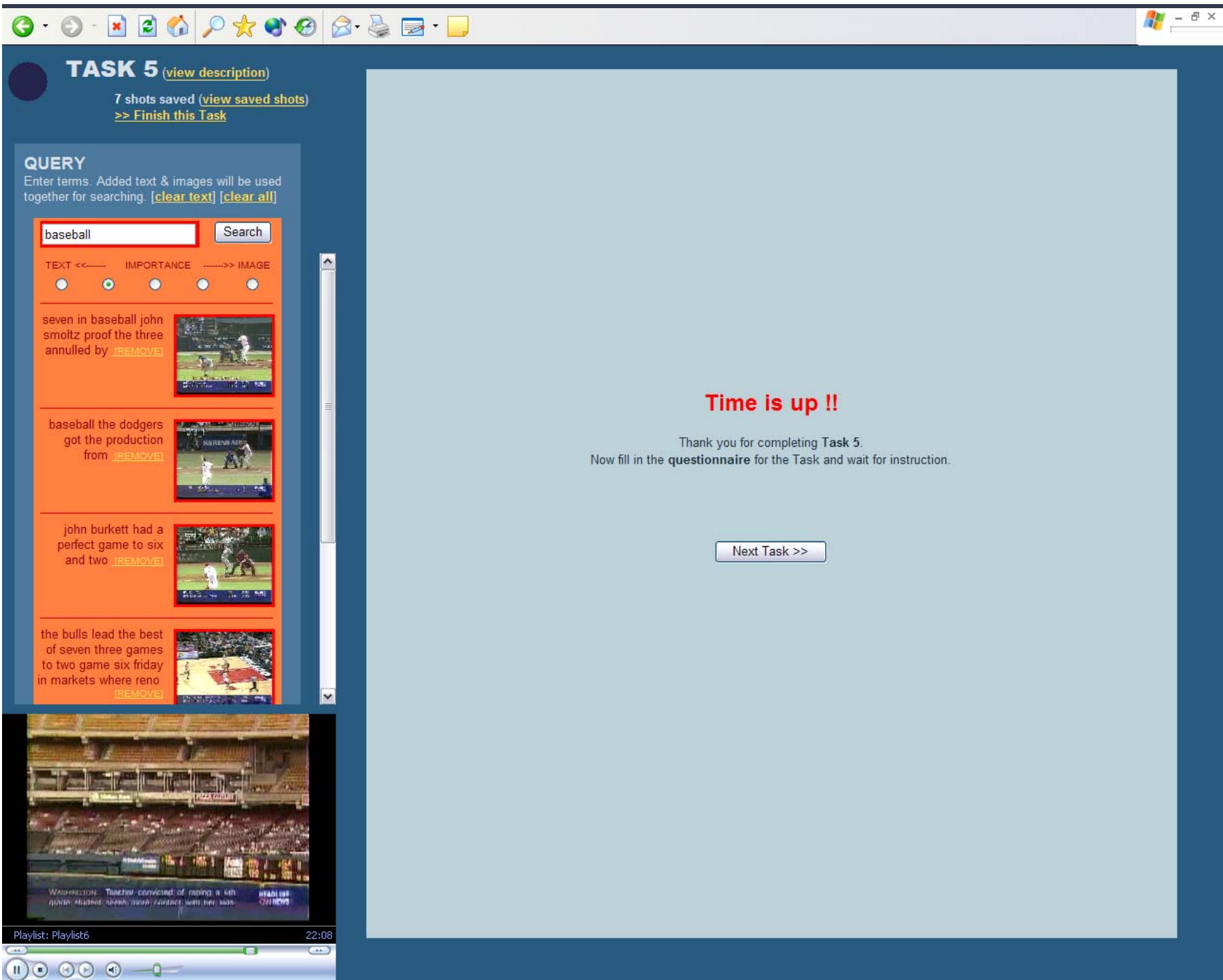






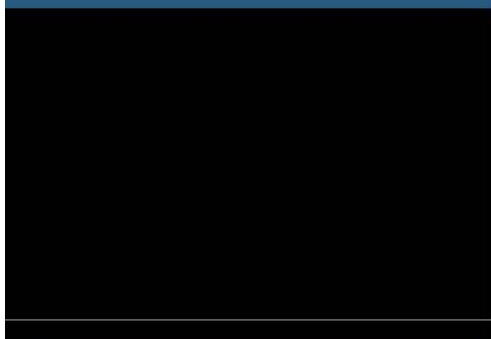




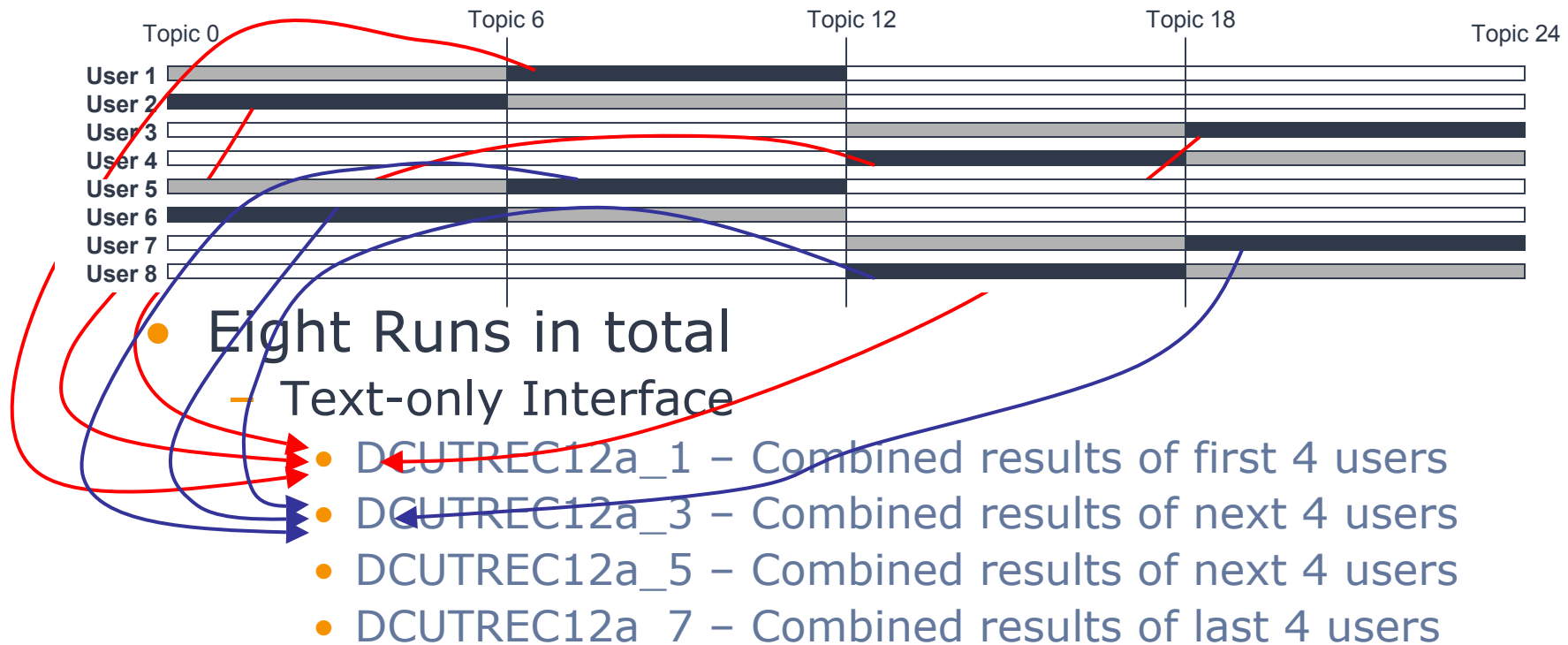


Demonstration

Text-only System Demonstration

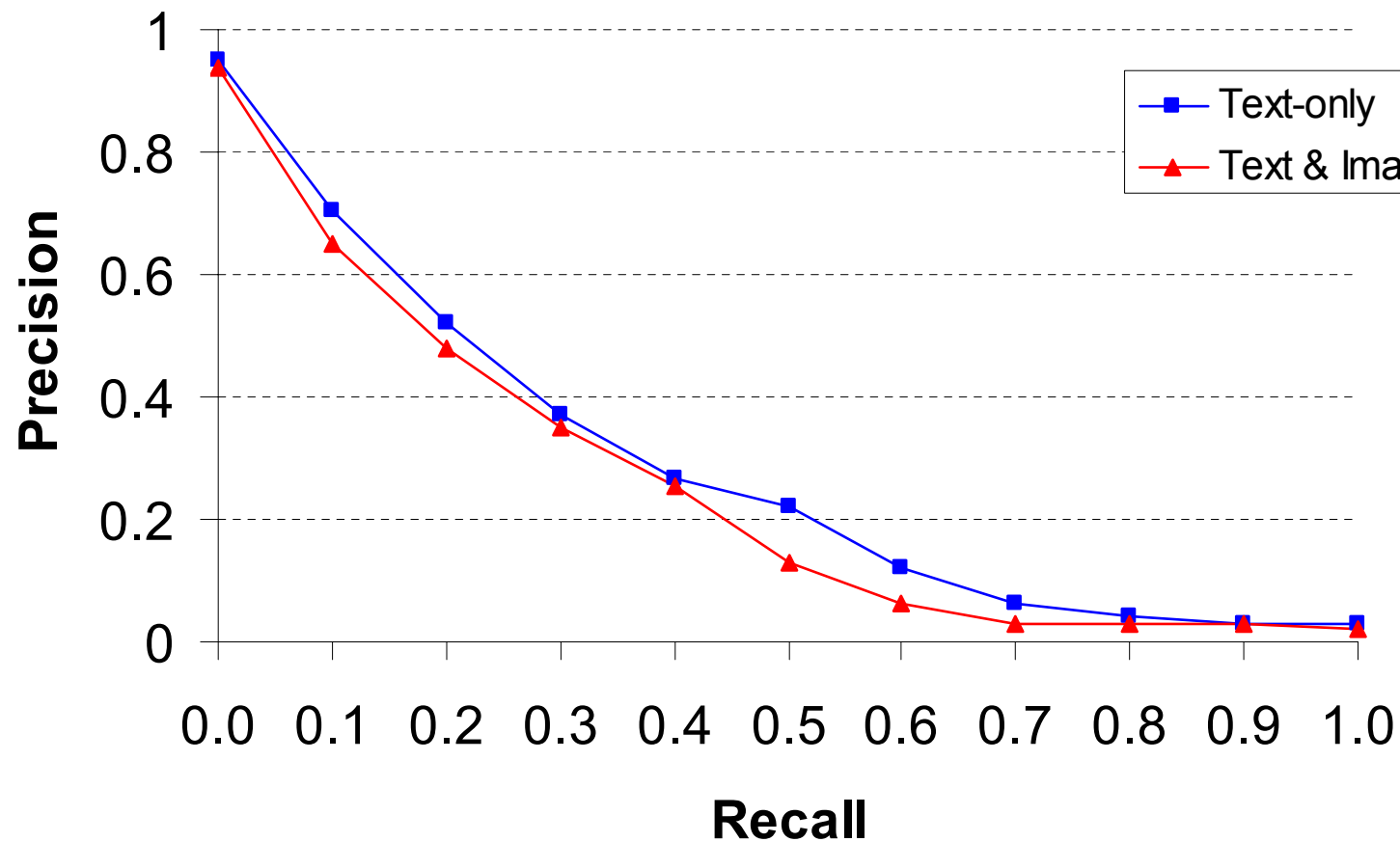


Submitted Runs

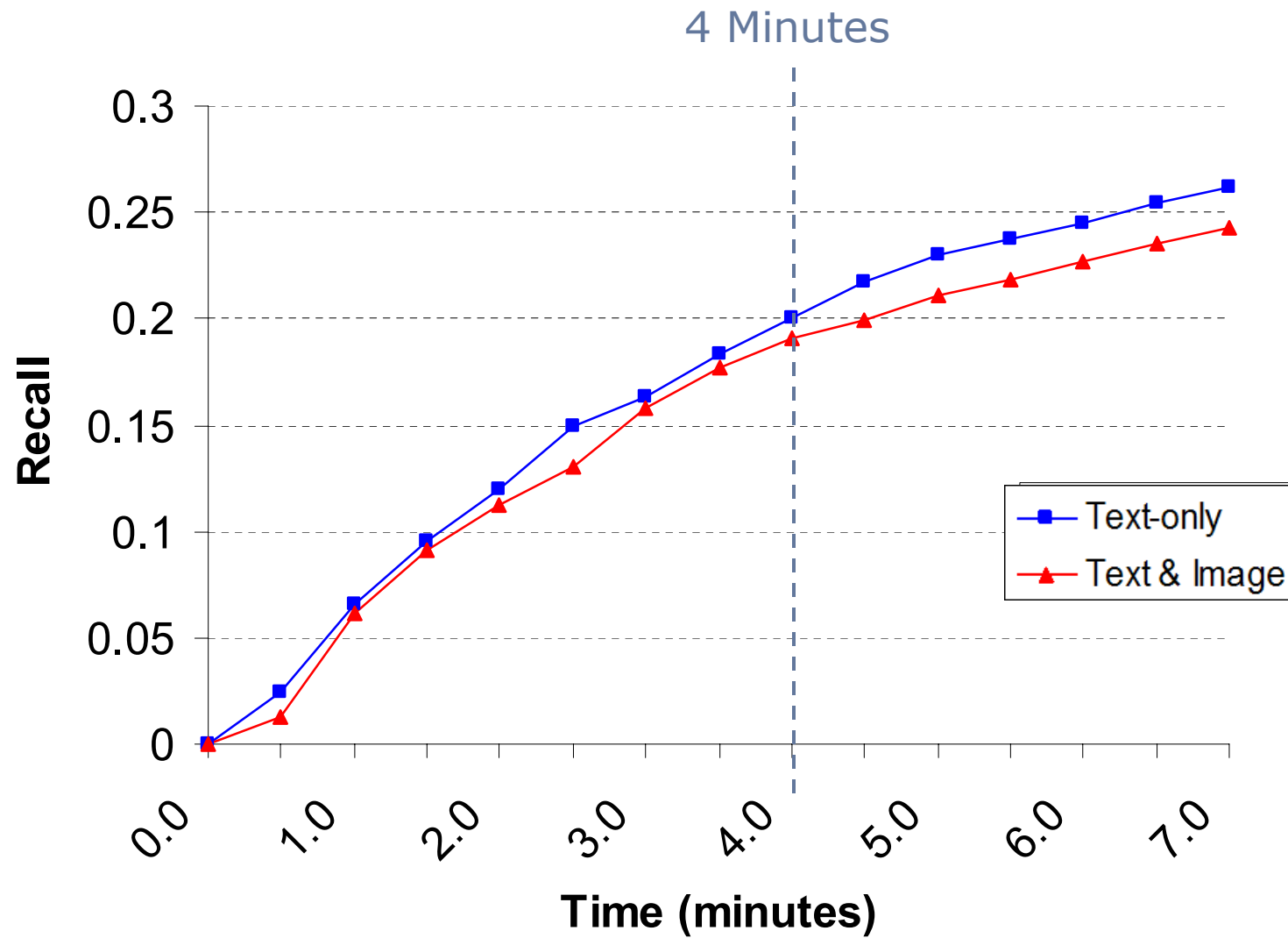


Precision Recall graph

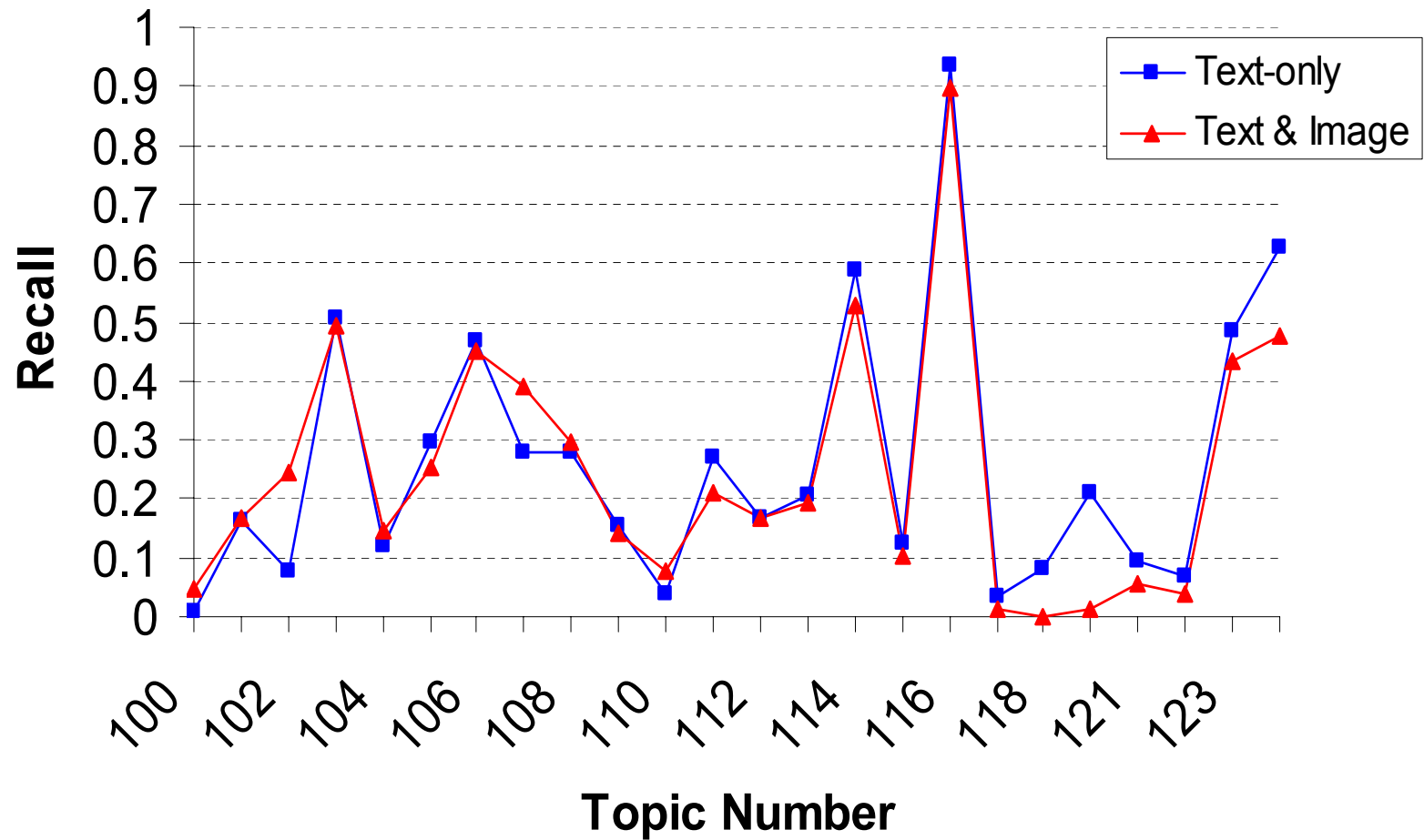
Aggregation of all 4 runs for each system



Examining time



Recall over Topic

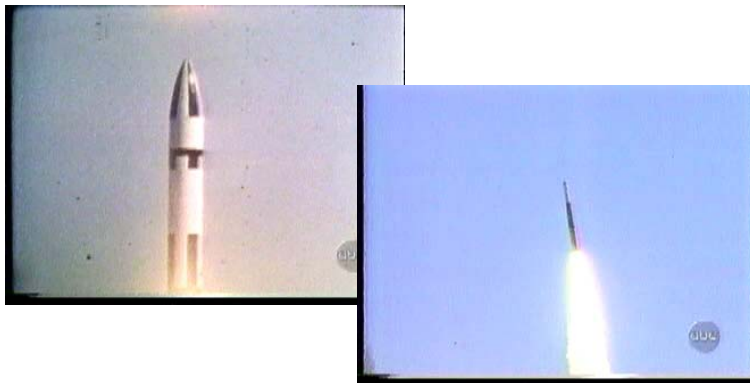


Text, Image & Feedback Queries



Topic 102

Find shots from behind the pitcher in a baseball game as he throws a ball that the batter swings at



Topic 107

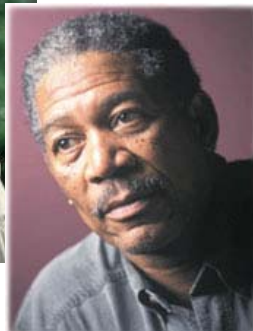
Find shots of a rocket or missile taking off. Simulations are acceptable

Text-only Queries



Topic 111

Find shots with a locomotive (and attached railroad cars if any) approaching the viewer

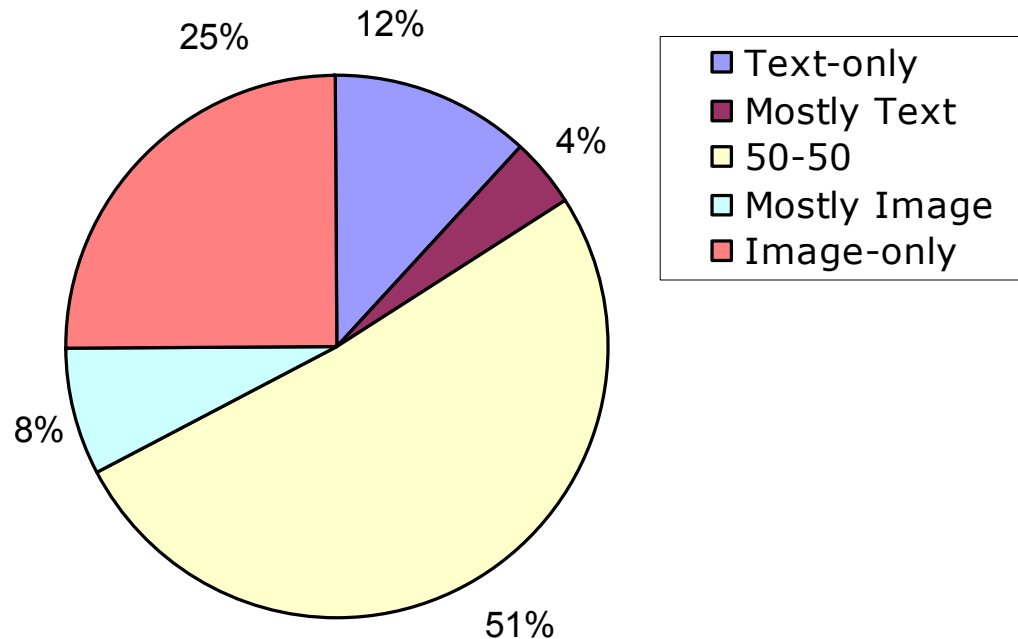
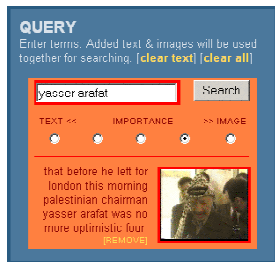


Topic 119

Find shots of Morgan Freeman

User Observations

- Average of 6 queries / topic (both systems)
 - 564 in total on the Text-only and 581 on Text, Image and Feedback
 - Of 581 Text, Image & Feedback Queries:
 - > 99% contain text and 81% contain an image
- When given the choice, users chose:



Conclusions

- Both systems perform comparably
 - Text-only seems to be slightly better than the text, image and feedback system
 - But not by any significant amount
- Why is this the case?
 - Text-only is better...
 - Users more comfortable with text querying...
 - Query response time of the text, image and feedback system was slower than text-only...
 - By a few seconds only over the seven minutes.
- We still have more work to do on evaluating the user data gathered during the experiments



Thank You