Interactive Search Using a Lexicon too Large to Remember

Marcel Worring, Bouke Huurnink,
Ork de Rooij, Cees Snoek, Andy Thean

Intelligent Systems Lab Amsterdam,
University of Amsterdam, The Netherlands

1TNO, The Netherlands
Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.

Find fragments such as this in the database
Generic Semantic Indexing

Semantic Pathfinder

Select Best of 3 Paths after Validation

Content Analysis Step

Style Analysis Step

Context Analysis Step
TRECVID 2004

Learned lexicon of 32 concepts

No concept is relevant
Learned lexicon of 101 concepts

Explosion might be relevant
Too many to remember
Should we bother?

TRECVID2005

Influence of Thesaurus Size on Video Search Performance

Mean Average Precision

Concept Detectors in Thesaurus

Yes size does matter.
So we need to help the user in selecting the right concept.
Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.

----------

Depends on what you are looking for
Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.
Suggestion based on text matching

Index concept descriptions
Represent as term vector
Only 363, so rather small collection

Need to increase recall?
Porter stemming algorithm
Character $n$-gramming, here sequences of 4 characters

We use the vector space model to match queries to descriptions
Pick detector that maximizes query/document similarity

Turns out that perfect match yields best performance
A wheeled vehicle that has two wheels and is moved by foot pedals.
Suggestion using ontology querying

“Find a report from the desert showing a house or car on fire.”

1. Identify objects in WordNet
Ontology querying

“Find a report from the desert showing a house or car on fire.”

2. Identify related concept detectors
Ontology querying

“Find a report from the desert showing a house or car on fire.”

3. Find most similar and specific detector using Resnik’s measure
Query: Find shots of the current president of America
Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.
Visual Similarity Space

Matrix containing all similarities between pairs of shots
Textual Similarity Space

Matrix containing all similarities between pairs of shots
Semantic probabilities

Concept Lexicon

<table>
<thead>
<tr>
<th></th>
<th>C₁</th>
<th>C₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>P₁</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₂</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Semantic Similarity Space

Use probabilities as a "feature vector"
Visual Thread Space

Define a shortest path through each cluster
Semantic Thread Space

We can do the same for semantics (and text)
Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.
The CrossBrowser
Other browsing dimensions

Time
The timeline of the original video

Keypoint based similarity
Requires interaction by the user, cannot be pre-computed

The concept for which the shot receives a top-rank
The RotorBrowser
Results

TRECVID 2006 Interactive Search Results

- 34 users of other video retrieval systems
- 2 users of MediaMill video search engine

Mean Average Precision

System Runs

0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
35

Additional images below graph showing visual results.
Results: CrossBrowser

Interactive Search Results

1. one or more emergency vehicles in motion
2. tall buildings and the top story visible
3. people leaving or entering a vehicle
4. soldiers escorting a prisoner
5. a daytime demonstration with one building visible
6. US Vice President Dick Cheney
7. Saddam Hussein with another persons face visible
8. multiple people in uniform and in formation
9. US President George W. Bush, Jr. walking
10. soldiers with weapons and military vehicles
11. water with boats or ships
12. people seated at a computer, display visible
13. people reading a newspaper
14. a natural scene; no buildings, roads, or vehicles
15. one or more helicopters in flight
16. something burning with flames visible
17. people dressed in suits, seated, and with flag
18. a person and at least 10 books
19. an adult person and a child
20. a greeting by kiss on the cheek
21. one or more smokestacks, with smoke
22. Condoleezza Rice
23. soccer goalposts
24. scenes with snow

Direct mapping to topic
Complex concept combination

Search Topic

Average Precision
Results: RotorBrowser

Interactive Search Results

1. one or more emergency vehicles in motion
2. tall buildings and the top story visible
3. people leaving or entering a vehicle
4. soldiers escorting a prisoner
5. a daytime demonstration with one building visible
6. US Vice President Dick Cheney
7. Saddam Hussein with another persons face visible
8. multiple people in uniform and in formation
9. US President George W. Bush, Jr. walking
10. soldiers with weapons and military vehicles
11. water with boats or ships
12. people seated at a computer, display visible
13. people reading a newspaper
14. a natural scene; no buildings, roads, or vehicles
15. one or more helicopters in flight
16. something burning with flames visible
17. people dressed in suits, seated, and with flag
18. a person and at least 10 books
19. an adult person and a child
20. a greeting by kiss on the cheek
21. one or more smokestacks, with smoke
22. Condoleezza Rice
23. soccer goalposts
24. scenes with snow

Average Precision
Our RotorBrowse run

Bookmarks and their judgement by NIST

Initial query and time dominant. (CrossBrowser would have been sufficient). Does help to find additional shots.
The set of threads to use is very dependent on the topic.
Decomposition: CrossBrowser

Time is a very important factor
Thread to use is not obvious
Lessons learned

Region based querying
For the current TRECVID topics of limited use

CrossBrowser versus RotorBrowser
For most topics initial query and time are contributing most to the final result, so CrossBrowser often sufficient
But in specific cases the use of additional threads can boost performance

The optimal threads
Do not exist, depends on the topic

Choosing the threads
For a novice user not evident from the visualization, performance of threads still too poor
Thanks for your attention