Outline

- System Overview
- Concept-Based Search
- Experiments & Results
- Conclusion
Outline

- System Overview
- Concept-Based Search
- Experiments & Results
- Conclusion
Automatic Search System

Framework

Multimedia Query

Find shots of water with boats or ships

Text-based Retrieval

Concept-based Retrieval

Visual-based Retrieval

Multi-Modal Fusion
Automatic Search System

- **Text-based search**
  - Keywords: expanded by WordNet
  - Transcript segmentation:
    shot-level, story-level, video-level
  - Result expansion for shot-level search:
    scores spread along the timeline
Automatic Search System

- Text-based search
- **Visual-based search**
  - Richer feature set
  - Feature selection & fixed-value fusion weight: MAP & consistency
    - 5 features involved
  - Several SVM classifiers for each feature
  - Weighted average multi-feature fusion
Automatic Search System

- Text-based search
- Visual-based search
- Concept-based search
  - Query-concept mapping
  - Text-concept mapping
  - Example-concept mapping
- More details come later.
Automatic Search System

- Text-based search
- Visual-based search
- Concept-based search

Fusion
- Weighted average
- Query-independent
Interactive Search System

- User interface
  - faster, faster and faster
  - Browsing functions

- Server end
  - Several options
Interactive Search System: UI

- Double-screen interface
- Multi-thread browsing
  - Temporal thread
  - Visual neighbor thread
- Frame-level browsing
- Browsing function
  - Forward, Backward, Bookmark
- Hotkey
Faster Browsing

Story browser
Frame-level browsing

Browsing

Rank list
Temporal thread
Visual neighbor thread
Labeling: Hotkey & Mouse
Refining Positive samples
Negative samples
Uncertain samples

Bookmark
Server end

- Distributed server end
- More options
  - 1 text-based server
  - 4 SVM models with different features
  - 2 concept-based servers
  - manually adjusted options Vs. default options
Outline

- System Overview
- Concept-Based Search
- Experiments & Results
- Conclusion
Concept-Based Search

- Well established approach
- Need theoretical guidance for practical issues
- Query-Concept Mapping (QUCOM)

**Example**

- tall building → Building, Cityscape, Conference_Building, ...
- scenes with snow → Snow

Query Image

- Boat/Ship, Waterscape, ...
Possible Solutions for QUOCOM

- User choice?
  - Text Match
    - ([Snoek, 2006], [Chang, 2006], et c)
    - Effective if well matched
    - Fails to consider
      - visual correlation
      - concept performance
      - concept distribution over the collection

- Concept Space
  - Search in Full Space
    - (e.g. SVM, KNN [Natsev, 2006], PMIWS [Zheng, 2006])
  - Search in Concept Subspace
Concept Selection via c-tf-idf Metric

**Concept Relevance Ranking**

<table>
<thead>
<tr>
<th>Definition in text area</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>tf</strong>: frequency of a term in a document → <em>term popularity</em></td>
</tr>
<tr>
<td>• <strong>idf</strong>: inverse document frequency of a term → <em>term specificity</em></td>
</tr>
</tbody>
</table>

**c-tf-idf: tf-idf for concept**

c: concept,  d: shot

\[
c - tf - idf(c, d) := freq(c, d) \log\left( \frac{N}{freq(c)} \right) = P(c \mid d) \log\left( \frac{N}{freq(c)} \right)
\]
**Insight of the *tf-idf* based Principle**

\[
c - tf - idf(c, d) := freq(c, d) \log\left(\frac{N}{freq(c)}\right)
\]

---

**c-tf-idf** is a good combination of query-dependent ranks and query-independent ranks, and a promising solution for QUCOM.
Two Settings for QUCOM

- **Automatic video retrieval (AVR)**
  - limited information as text input, and possibly, image examples

- **Interactive video retrieval (IVR)**
  - unrealistic to ask user provide relevant concepts
  - Infer the implicit semantic concepts by explicit user feedback

- **QUCOM should be**
  - On a per query analysis basis, on-the-fly,
  - Combat against varied concept detection performance
  - Scalable to
    - Concepts in a given lexicon
    - Video archive size
Concept-Based Search: Search

- Search in concept subspace
- Impact of dimension of subspace

Experiment on TRECVID 2006, interactive search

Experiment on TRECVID 2006, automatic search
Inferring implicit concepts through explicit feedback: Interactive Search

Interactive Search

- Using relevance feedback as examples
- Higher efficiency: Vs. user-provided examples
  - Pre-computed offline
- Lower user labor: Vs. manual concept selection
- Better performance: Vs. previous system
  - 65% improvement upon previous method (without using concepts)
  - Experiment on TRECVID 2006, interactive search
Concept-Based Search: Lexicons

- **LSCOM-lite**
  - 39 concept detectors from HLF task
- **LSCOM**
  - 374 concepts chosen from LSCOM
- **Impact of quality & quantity?**
  - Experiment on TRECVID 2006, interactive search
Outline

- System Overview
- Concept-Based Search
- Experiments & Results
- Conclusion
Automatic runs

- Run1: text : 0.011
- Run2: image + LSCOM-Lite : 0.042
- Run3: text+image : 0.038
- Run4: text+image+LSCOM : 0.043
Run1

- Text-based search
  - Helpful to topics about Object
  - Useless to topics about Event or Scene

- Unsatisfactory upon non-news video
Run2, Run3, Run4

Run2 Vs. Run 4
- Concept detectors from LSCOM (except 39 concepts from HLF) are trained upon different dataset.

Run2 Vs. Run3
- Involving concept-based search brings improvement.

MAP

<table>
<thead>
<tr>
<th>Run1: text</th>
<th>Run2: image + LSCOM lite</th>
<th>Run3: text + image</th>
<th>Run4: text + image + LSCOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0104</td>
<td>0.041</td>
<td>0.0376</td>
<td>0.0426</td>
</tr>
</tbody>
</table>
Interactive runs

- Run5: expert with manually adjusted options : 0.209
- Run6: expert with default options : 0.171
- RunS: novice with default options : 0.149
Outline

- System Overview
- Concept-Based Search
- Experiments & Results
- Conclusion
Conclusion

- Concept-based search is fruitful and complement to text and visual search
- A easy-to-use UI is essential to interactive search
- User can make-up the drop in automatic