MCG-ICT-CAS TRECVID 2008 Automatic Video Retrieval System

Juan Cao, Yong-dong Zhang,
Bai-lan Feng, Xiu-feng Hua, Lei Bao, Xu Zhang

Multimedia Computing Group
Institute of Computing Technology
Chinese Academy of Sciences

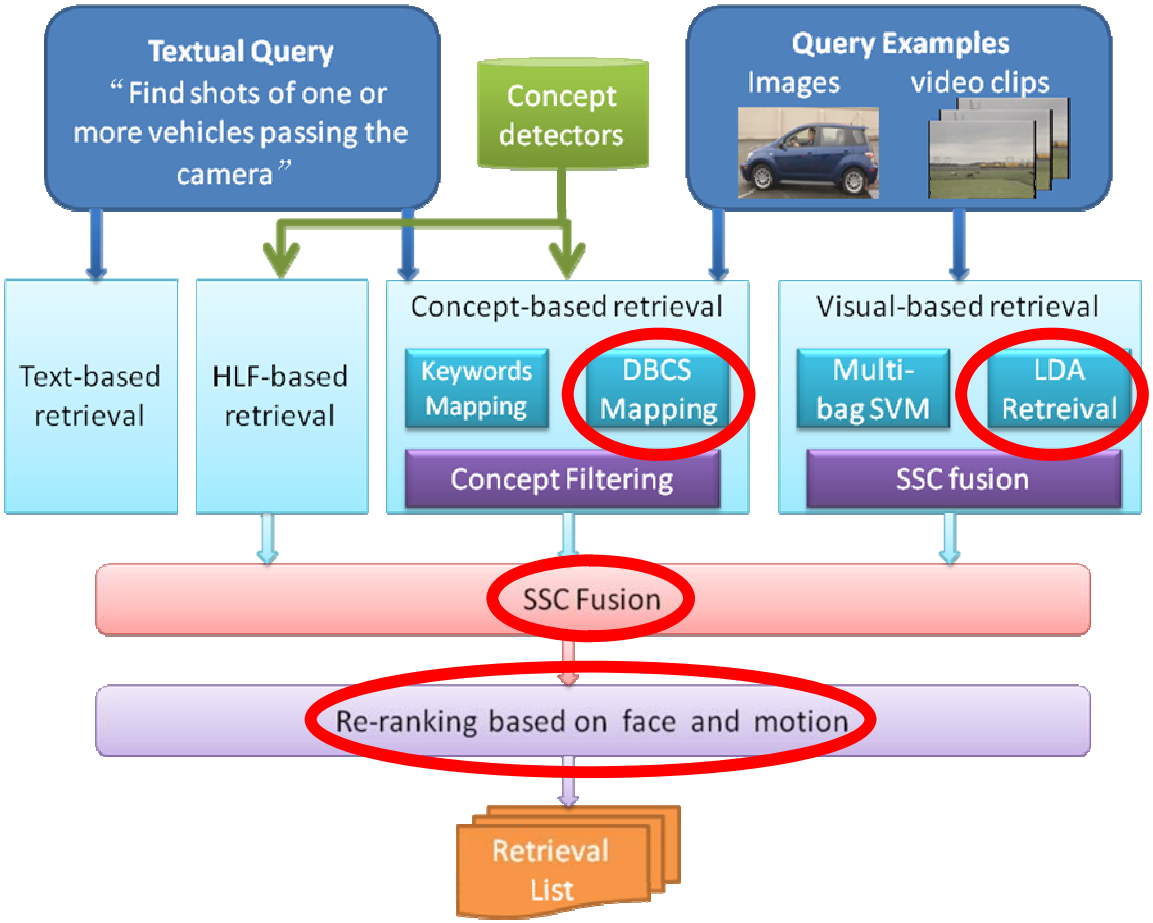
INSTITUTE OF COMPUTING TECHNOLOGY 中科院计算所

NIST TRECVID Workshop
November 17,2008

Outline

- **Overall system**
- Review of baseline retrieval
- Performance analysis
  - Concept-based retrieval
  - Re-ranking
  - Dynamic fusion
- Conclusion

System Overview



Review of baseline retrieval

INSTITUTE OF COMPUTING TECHNOLOGY

- **Text-based retrieval** **0.009**
  - ASR shot matching

    A window of 3 shots
  - Pre-processing

    Stop words removing stemming
  - Indexing

    lucence

Review of baseline retrieval

INSTITUTE OF COMPUTING TECHNOLOGY

- Text-based retrieval **0.009**
- **Visual-based retrieval** **0.033**
  - Feature extraction

    EH CM Sift-visual-keywords

    Early fusion and **LDA embedding**
  - Retrieval model

    Multi-bag SVM cosine-similarity
  - Fusion

    SSC dynamic fusion

Review of baseline retrieval

INSTITUTE OF COMPUTING TECHNOLOGY

- Text-based retrieval 0.009
- Visual-based retrieval 0.033
- **HLF-based retrieval** 0.029
  - Concept detectors
    - CU-VIREO374
  - Retrieval Model
    - Multi-bag svm

[Acknowledgement]: Thank Dr. Yu-Gang Jiang for great help in the experiments.

Review of baseline retrieval

INSTITUTE OF COMPUTING TECHNOLOGY

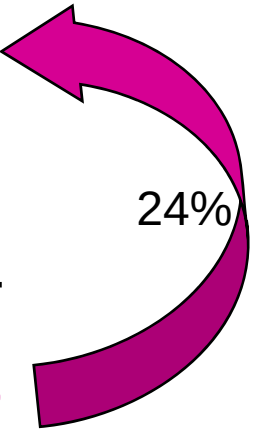
- Text-based retrieval 0.009
- Visual-based retrieval 0.033
- HLF-based retrieval 0.029

- **Concept-based retrieval 0.044**
  - Keywords mapping
  - DBCS mapping

Review of baseline retrieval

INSTITUTE OF COMPUTING TECHNOLOGY

- Text-based retrieval 0.009
- Visual-based retrieval 0.033
- HLF-based retrieval 0.029



- **Concept-based retrieval** 0.044
- **Re-ranking** **0.036**
  - Face
  - motion

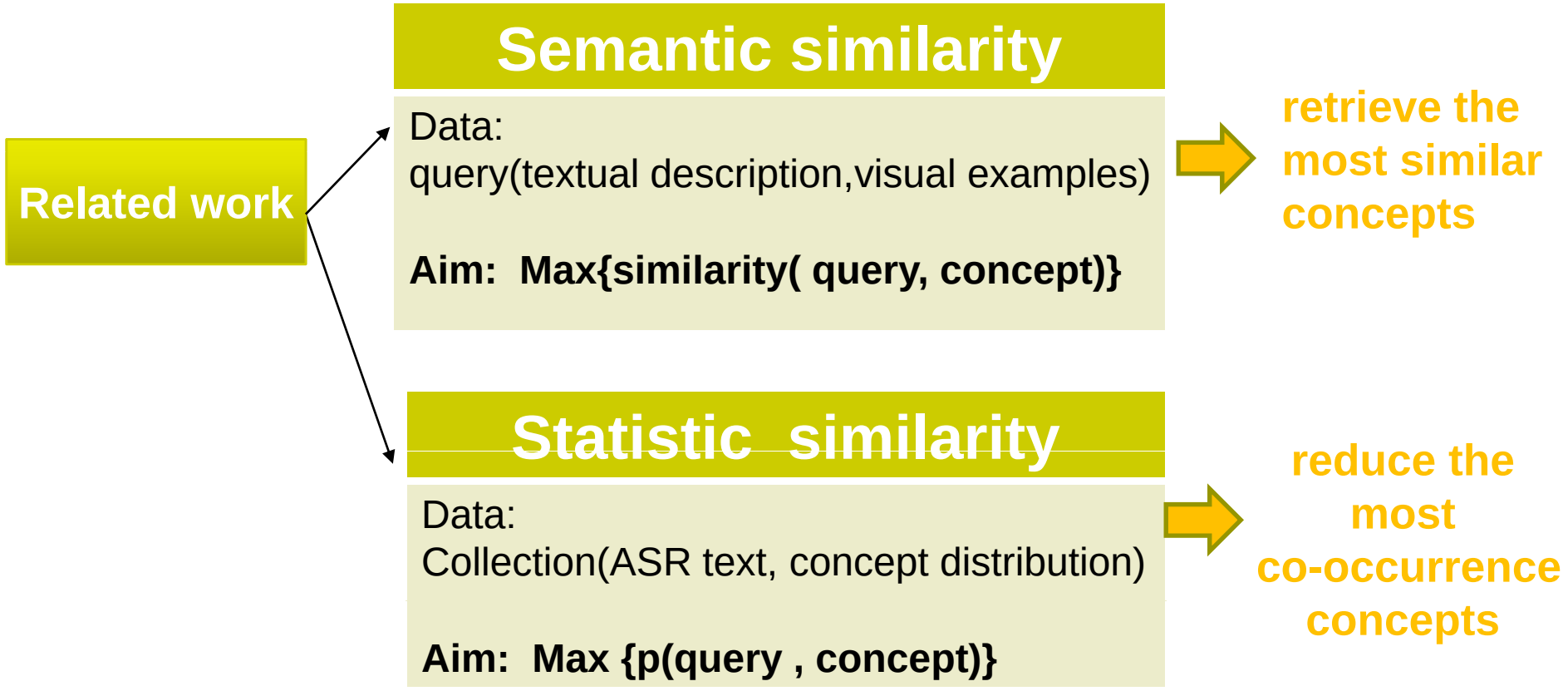
Review of baseline retrieval

- Text-based retrieval 0.009
- Visual-based retrieval 0.033
- HLF-based retrieval 0.029

- **Concept-based retrieval** 0.044
- **Re-ranking** **0.036**
- **SSC Dynamic fusion**

Query-to-concept mapping

INSTITUTE OF COMPUTING TECHNOLOGY



Query-to-concept mapping

What is useful ?

- Discriminability-ranking
  - The distributions fluctuate widely between the given category and the others, but remain stable within this one.
- Factors
  - **Difference** of the concept distribution
  - Detector performance
  - Collection characteristic

Distribution Based Concept Selection Framework(DBCS)

INSTITUTE OF COMPUTING TECHNOLOGY

VAC: the difference between categories

$$VAC(t,c_i) \leftarrow \sum_{j \neq i} sign(F(t,c_i) - F(t,c_j))(F(t,c_i) - F(t,c_j))^2$$

VIC: the difference within the given category

$$VIC(t,c_i) \leftarrow \frac{1}{n_i} \sum_{s \in c_i} (F(t,s) - F(t,c_i))^2$$

Discriminability-score

$$Score(t) = VAC(t,c_i) / VIC(t,c_i)$$

- Where ***F(t,s)*** is the distribution function of concept ***t*** in shot s.

Example-1
Discriminability-similarity consistency

INSTITUTE OF COMPUTING TECHNOLOGY

- Topic248 Find shots of a crowd of people , outdoors , filling more than half of the frame area

DBCS approach infAP=0.321	
Crowd	1.40
People_Marching	0.92
Demonstration_Or_ Protest	0.64
Protesters	0.55
Dark-skinned_People	0.52

Text selection approach infAP=0.203
Crowd
Outdoors
Person

Factor-1: outdoors and person also frequently occur in other case.

Factor-3: collection characteristic

Example-2
Discriminability-similarity inconsistency

INSTITUTE OF COMPUTING TECHNOLOGY

- Topic261 Find shots of one or more people at a table or desk , with a computer visible

DBCS approach infAP=0.116	
Attached_Body_Parts	0.55
Classroom	0.30
Medical_Personnel	0.27
Body_Parts	0.25
Hand	0.23

Text selection approach infAP=0.012
Computer
Computer_Or_Television_Screens
person

Factor 2: computer detector is not reliable

Re-ranking

- face and motion factors
  - shot-level **average** face size and position
  - shot-leve **principal** motion direction and intensity

$$Score' = Score + FactorScore \times FactorCoefficient$$

➢Shot-level vs. Keyframe-level
Extract the stable factor

➢Re-ranking selection
Reduce the negative effect

Dynamic fusion

- **Smoothed Similarity Cluster(SSC)**
  - A feature undergoes a rapid change in its normalized scores is likely to perform better than a feature which undergoes a more gradual transition.

$$SC = \frac{\frac{1}{1000}\sum_{n=1}^{1000}(score(n) - score(n+1))}{\frac{1}{N}\sum_{n=1}^{N}(score(n) - score(n+1))}$$

[P. Wilkins,2007]

$$SSC = \frac{median(SC)}{standard\ \ deviation(SC)}$$

SC is unstable in real noisy data.

$$Run\ \ Weight = \frac{Run\ \ SSC\ \ Score}{\sum All\ \ SSC\ \ Scores}$$

In our system, all fusion processes are realized by SSC method.

Outline

- Overall System
- Review of baseline retrieval
- Performance analysis
  - Concept-based retrieval
  - Re-ranking
  - Dynamic fusion
- **conclusion**

INSTITUTE OF COMPUTING TECHNOLOGY

Overview of submitted and unsubmitted runs

INSTITUTE OF COMPUTING TECHNOLOGY

Run Description	Mean InfAP
Run 1: Text baseline	0.009
Run2*: Visual baseline(Multi-bag SVM)	0.024
Run3*: Visual baseline(LDA)	0.028
Run4: SSC(Run2, Run3)	0.033
Run 5: HLF baseline(svm, CU-VIREO374)	0.029
Run 6: HLF baseline +re-ranking	0.036
Run 7*: Concept retrieval(text map, CU-VIREO374)	0.026
Run 8*: Concept retrieval(DBCS map, CU-VIREO374)	0.039
Run 9*: SSC(Run7 + Run8)	0.043
Run 10: SSC(Run5 + Run9)	0.053
Run 11: SSC(Run4 + Run9)	0.067

* is the unsubmiited run

Overall performance analysis

INSTITUTE OF COMPUTING TECHNOLOGY

Automatic search runs of TRECVID2008

Conclusion-1

INSTITUTE OF COMPUTING TECHNOLOGY

- Concept-based retrieval is a promising direction.

- DBCS mapping method can achieve a **stable good performance.**
  - **The difference** of the concept distribution is more useful than the distribution itself .
  - Select concepts independent of the detector performance is not reasonable.

Conclusion-2

INSTITUTE OF COMPUTING TECHNOLOGY

- Face and motion based re-ranking technology is important for some special topics.
  - Shot-level feature is stable
  - Reducing the negative effect is important

Conclusion-3

- SSC dynamic fusion can make improvement in more than 80% cases, especially in the case of fusing different features.

Thank you!

Any more details can contact:

caojuan@ict.ac.cn