Supervised Reranking for Web Image Search

Linjun Yang[1] and Alan Hanjalic[2]

[1] Microsoft Research Asia, China
[2] Delft University of Technology, The Netherlands

Current Web Image Search Ranking

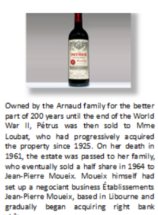
Query:

Red Wine

- The images are indexed mainly using the text in the associated web pages
- Feature-based ranking model

Problem Cases (1)

Example Query: Red Wine



(a) Missing

(b) Mismatch

Problem Cases (2)

Example Query: George W. Bush

(a) Mismatch between the image content and the words

(b) These images, though all relevant, are hard to differentiate and rank based on text

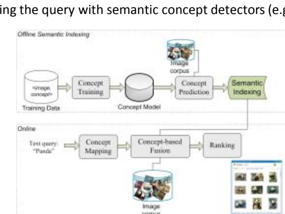
Content-aware Image Search

- Problem Summary
  - The associated text is not the essential description of images, and tends to be mismatching.
- Existing Solutions
  - Concept-based image search
  - Visual reranking

Concept-based Image Search

(a.k.a Semantic Indexing)

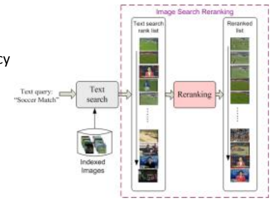
Linking the query with semantic concept detectors (e.g. LSCOM)



\+ Bridging the semantic gap
\- Scalability: it requires a large amount of labeled images to train concept models
\- Index capability: indexing images as concepts leads to information loss

Visual Reranking

- Reorder the text-based ranking list with visual consistency
  - Based on initial text-based search results
  - Exploit the intrinsic consistency



+ Lightweight, unsupervised, and scaled-up
- Hard to trade-off the contribution from the textual and visual domains
- The reranking model is hand-designed, which may not capture users' true needs

Comparison of Existing Approaches

	Concept-based Search [1]	Visual reranking [2,3]
Scalability	Challenging	Easy to scale up
Index capability	Weak (concept)	Good (visual feature)
Model representability	Good	Weak

[1] C. G. M. Snoek and M. Worring, "Concept-based video retrieval," Foundations and Trends in Information Retrieval, 4(2):215–322, 2009.
[2] X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, X.-S. Hua, "Bayesian video search reranking," ACM Multimedia 2008.
[3] W. H. Hsu, L. S. Kennedy, and S. Chang, "Reranking Methods for Visual Search," IEEE Multimedia 14, 3 (Jul. 2007), 14-22.

Supervised Reranking

Comparison of Methods

	Concept-based Search [1]	Visual reranking [2,3]	Supervised reranking
Scalability	Challenging	Easy to scale up	Easy to scale up
Index capability	Weak (concept)	Good (visual feature)	Good (visual feature)
Model representability	Good	Weak	Good

[1] C. G. M. Snoek and M. Worring, "Concept-based video retrieval," Foundations and Trends in Information Retrieval, 4(2):215–322, 2009.
[2] X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, X.-S. Hua, "Bayesian video search reranking," ACM Multimedia 2008.
[3] W. H. Hsu, L. S. Kennedy, and S. Chang, "Reranking Methods for Visual Search," IEEE Multimedia 14, 3 (Jul. 2007), 14-22.

Supervised Reranking

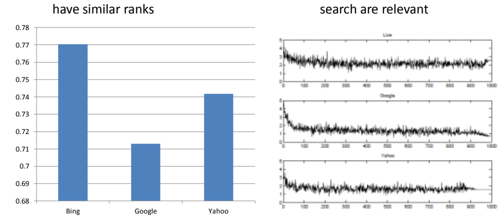
- Challenging problem
  - How to extract ranking features for a pair of textual query and visual document?
- Solution
  - Reranking features
    - Textual domain
      - Based on the initial text-based ranking
    - Visual domain
      - the textual query as the ranked images in the initial result
  - Learning
    - Adapted Ranking SVM

$$\min \frac{1}{2}((\boxed{\frac{w_0}{\alpha}})^2 + \sum_{t=1}^{Z} w_t^2) + C\sum \xi_{jk}^{i}$$

Trade-off between contribution from the textual and visual domains

Feature Design Principles

- Visual consistency
  - Visually similar images should have similar ranks
- Pseudo Relevance feedback
  - Top ranked images in the text-based search are relevant



- Data set
  - collected top 1000 images with 29 queries from 3 image search engines

Reranking Features

- Visual consistency
  - Nearest neighborhood
  - Reverse neighborhood [4]
- Pseudo-relevance feedback
  - Models from top results
- Initial ranking based feature

HV_N	Hard Voting of Neighbors
RSV_N	Initial Rank based Soft Voting of Neighbors
$NRSV_N$	Neighbor Rank Weighted Initial Rank based Soft Voting of Neighbors
HV_R	Hard Voting of Reciprocal Neighbors
RSV_R	Initial Rank based Soft Voting of Reciprocal Neighbors
NSV_R	Neighbor Rank based Soft Voting of Reciprocal Neighbors
$NRSV_R$	Neighbor Rank Weighted Initial Rank based Soft Voting of Reciprocal Neighbors
PRF_d	Local Density Estimation for PRF
PRF_{dv}	Duplicate Voting for PRF
PRF_{sdv}	Soft Duplicate Voting for PRF
IR	Initial Ranking

[4]. R. H. van Leuken, L. Garcia, X. Olivares, and R. van Zwol, "Visual diversification of image search results," WWW 2009.

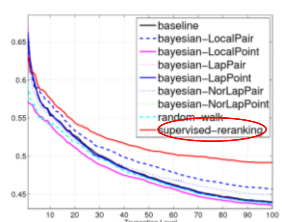
Think about it Broadly…

- The origin of the scalability
  - Embedding the queries into features
    - Feature extracted from <query, document> pairs
  - The model to be learned is independent of the queries
- Related schemes
  - Learning to rank
    - Features extracted from <keyword, text document> pair
      - *tf-idf*, BM25, LMIR, etc.
  - Distance metric learning
    - Features extracted from <image, image> pair

$$x^t A y = \sum_{i,j} a_{ij} x_i y_j$$

  - Query-relative Learning [Krapac et al., CVPR'10]

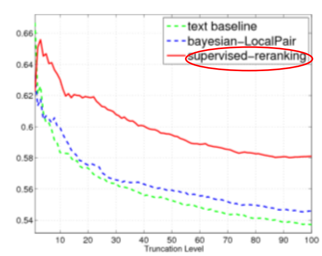
Performance Comparison (1)



- Data set
  - collected top 1000 images with 29 queries from 3 image search engines
- Evaluation measure
  - NDCG

- Supervised reranking outperforms the other unsupervised ones and text baseline
- The superiority of supervised reranking over Bayesian reranking is *statistically significant*

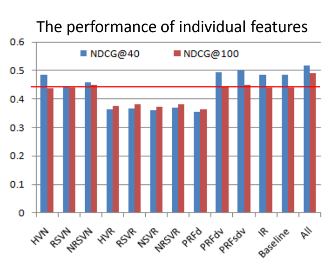
Performance Comparison (2)



- MSRA-MM dataset[1]
  - 68 queries, 1000 top images, from Bing
- Evaluation measure
  - NDCG

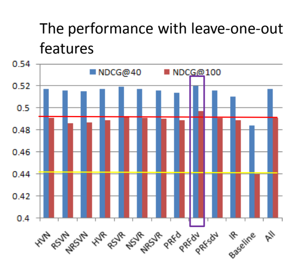
[1] http://research.microsoft.com/en-us/projects/msrammdata/

Individual Performance of Features

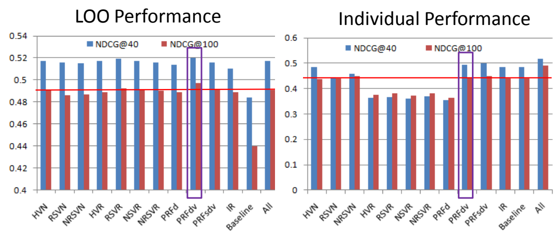


- The features have reasonable performance on their own.

Leave One Feature Out

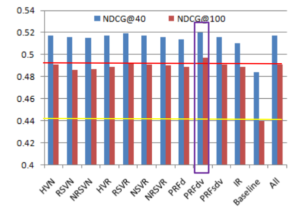


- Leaving some feature out improves the performance
  - Due to the correlation among features, good features may not contribute positively



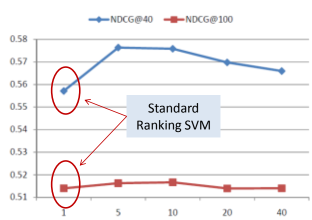
Leave One Feature Out

The performance with leave-one-out features



- Leaving some feature out improves the performance
  - Due to the correlation among features, good features may not contribute positively
  - It suggests automatic feature selection

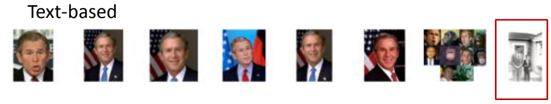
Adapted Ranking SVM



- Adapted Ranking SVM performs better than standard Ranking SVM (alpha=1).
- Setting alpha = # of visual features achieves the best result.

Reranking Results (query: George W. Bush)

Text-based



Reranking

Reranking Results (query: George W. Bush)

Text-based

Reranking

Conclusion and Future Work

- Conclusion
  - We proposed a new scheme, *supervised reranking*, for web image search.
  - Supervised reranking achieves both effectiveness and scalability by learning a query-independent model.
- Future work
  - Automatic feature selection for supervised reranking.
  - Supervised ranking with textual and visual ranking features.

Thanks !