Multimedia information retrieval beyond ranking

Michel Crucianu

http://cedric.cnam.fr/~crucianm/

CEDRIC lab, Vertigo research team

Conservatoire National des Arts et Métiers, Paris

Outline

- User needs
- Directions: complex answers, complex queries
- Complex answers
  - Step 1: clusters, maps, complementary clusters
  - Step 2: map to a known structure
  - Step 3: find and return a knowledge structure
- Complex queries
  - Step 1: multimedia concept detectors as queries
  - Step 2: structures as queries
- Conclusion

User needs?

Work on user needs in Multimedia IR mostly explored intent (the 'why': "reason, purpose or immediate goal")

1. Usually by extracting information from search engine log files
2. Sometimes based on interviews with professionals
3. More recently, using social-Web mining [HKL12]

but

- All approaches are biased by current abilities of search engines
  - Novel tools let users imagine new ways of using them + new needs to satisfy
- User need ~ 'what' + 'why'?
  - Unexpressed vs. expressed need, long term vs. immediate need…

Intent categories

Image search [LKM10]

1. Knowledge organization: gain knowledge from image(s)
2. Mental image: users have particular image content in mind
3. Navigation: find image known to exist without knowing its content
4. Transaction: find image to (re-)use as an object

Video search [HKL12]

1. Information: obtain knowledge by watching the video
2. Experience – learning: acquire skills by experience
3. Experience – exposure: have particular experiences
4. Affect: video as an immediate means to change mood, entertain
5. Object: video as an object for subsequent use

User needs

What studies of user intent tell us about user needs

- Partly dependent on the nature of the content (text, video…)
- Not mutually exclusive, often mixed and underexpressed
- An individual content item is the final goal in a minority of cases
- Even then, search sessions often have many steps

Conclusions regarding user needs

- Safe to respond to a superset of the needs estimated from the user query
- The underlying organization of the results helps users
  - → drive the search session
  - → identify the appropriate answers to their multiple needs
  - → get a more global understanding of what can be relevant to their query

Search results ranking

Early introduction in IR, inherited in Multimedia IR

Principle

- Define a score ('relevance') based on
  - How similar a candidate answer is to the query
  - How important / popular / authoritative the candidate is
- Return a list of answers ranked by decreasing score

Assumptions

- Users have teletype-like interfaces, so all relevance information should be squeezed in a one-dimensional score
- A user is looking for one or several content items
- A user might even be satisfied with the top-ranked content item
- What matters is the topic (the 'what')

Limitations of ranking

Related to the relevance score

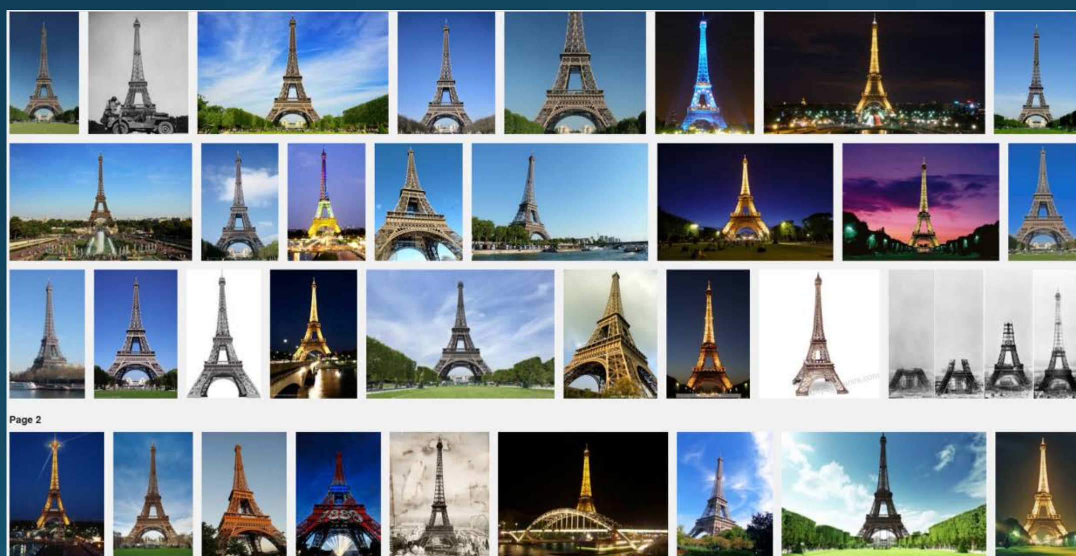
- Focus is only on 'what' and on 'popularity'

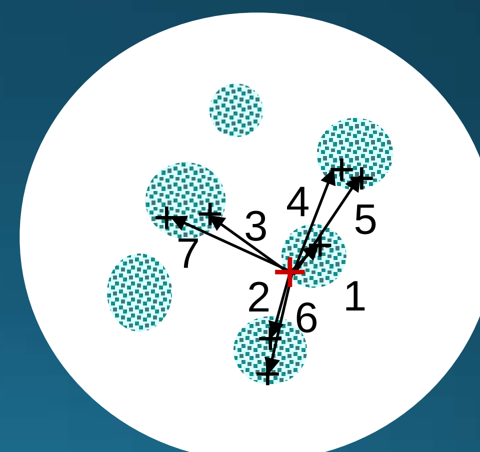
Related to the ranking of results

- Many results can be found 'relevant', but the user will seldom look beyond the top k (k << 1000) → low recall
- No information regarding
  - Different interpretations of the query (but introduction of auto-completion)
  - Different possible 'views' over the results
  - An underlying organization of the results
- No way to obtain a 'digest' of the results
  - Unless some 'authority' has a well-prepared and top-ranked answer… not automatically updated when available content evolves!

Limitations of ranking (images)

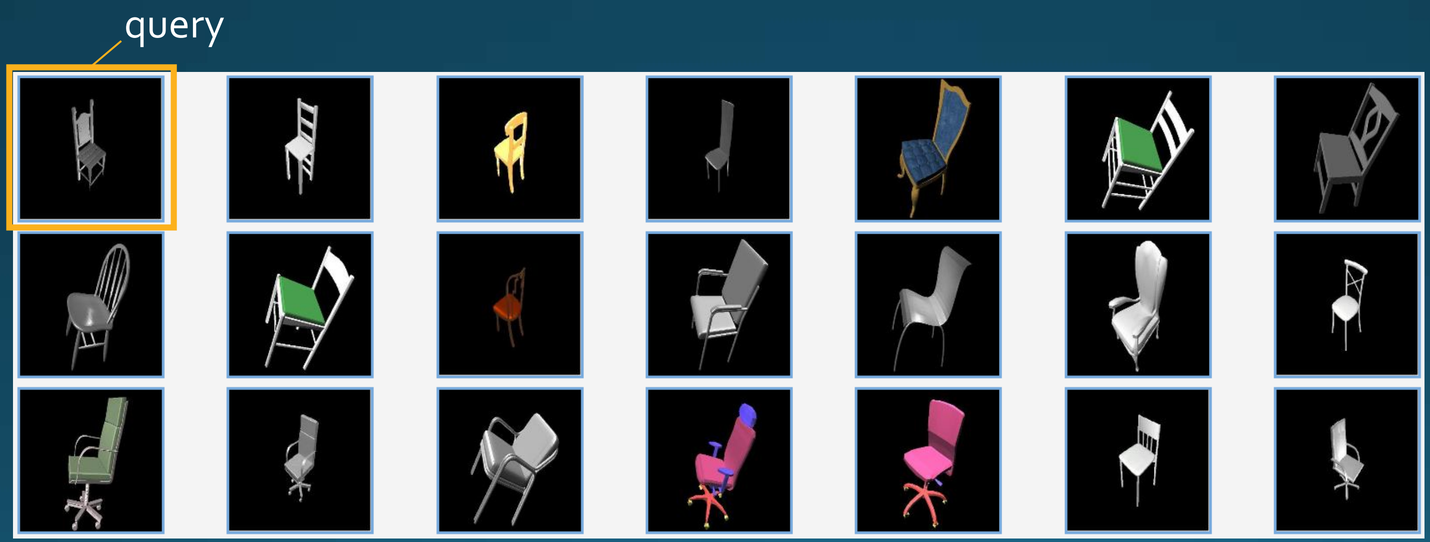
Keyword-based retrieval of images; results sorted by decreasing ‘relevance’, top ↘ bottom and left ↘ right





Limitations of ranking (3D obj.)

Content-based retrieval of 3D objects; results sorted by decreasing similarity to query



Improvements to ranking

Recent work attempts to remove some of the limitations

- Diversity: sample several clusters in the top *k* results
- Intent: identify and consider the 'why', not just the 'what'
- 'Slow search': more time to provide more informative answers

or go beyond ranking by

1. Collecting more information regarding the results
2. Displaying this information to the user

→ Answers should be more complex than just ranked lists

→ Opportunity to also ask more complex queries!

Beyond ranking:

complex answers

Result clustering

Principle: cluster relevant results, then return clusters (possibly ranked)

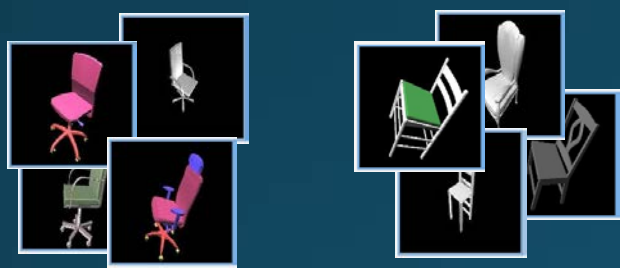
Result clustering

→ Highlights some (basic) structure in the results

But doesn’t show inter-cluster relations

Complementary clusterings

Goal: find complementary clustering criteria for the results, then return the clustering according to each criterion



[*with / without wheels*]

[*with / without armrests*]

→ Allows to identify complementary criteria / facets / views to structure the set of results

Complementary clusterings (2)

Principle

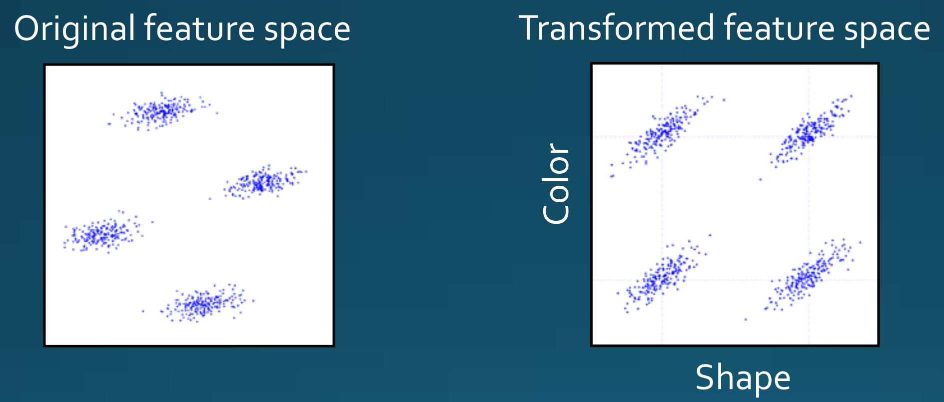
- Several alternative clusterings are obtained for the same dataset
- Vector representations: each clustering performed in specific, arbitrarily oriented subspace
- Complementarity: clustering in one subspace provides little or no information regarding clustering in the other subspace(s)

Methods

- Quite diverse methods were proposed (e.g. [DB14], [PC10])
- [DB14]: sequential discovery of subspaces where data clusters in a different way than in already found subspaces

Complementary clusterings (3)

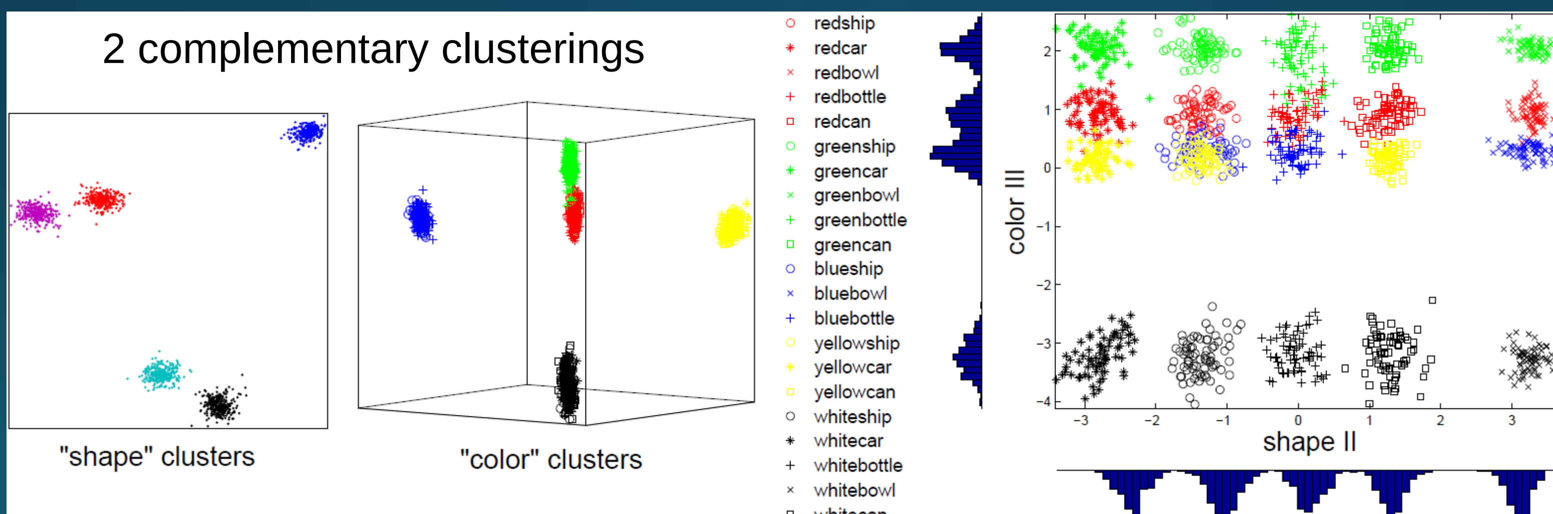
- [PC10]: inspired by Tree-Component Analysis, but mutual information is computed between clusters in different subspaces



Remaining challenge: give 'meaning' to the clustering criteria and to the corresponding clusters (e.g. by finding appropriate 'labels')

Illustrative results [PC10]

Data (from COIL-100):
21 object classes,
72 viewpoints/class
Global descriptions



Maps

Goal: display specific / selected or more global relations between answers to a query

→ Provide more insight into how and why some answers are relevant to a query

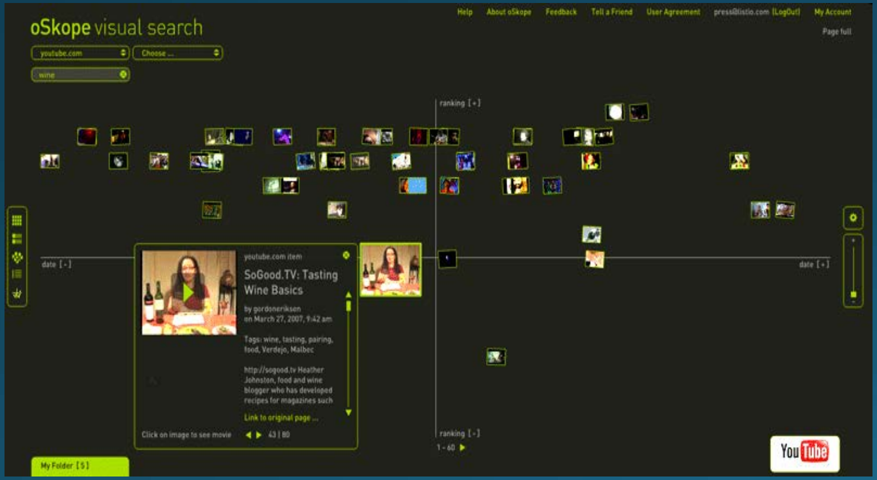
- Display relations between content items (or groups)
- Interpretable dimensions: the positions of content items along these dimensions are meaningful

Maps (2)

Approaches

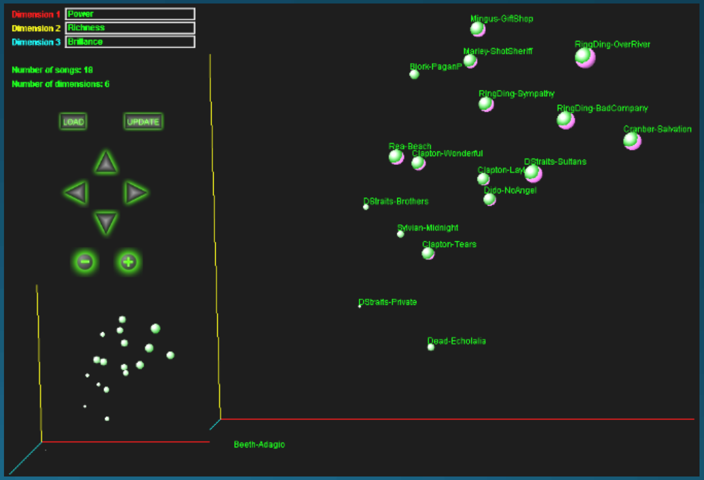
1. Prior selection of interpretable dimensions

Retrieval score & date



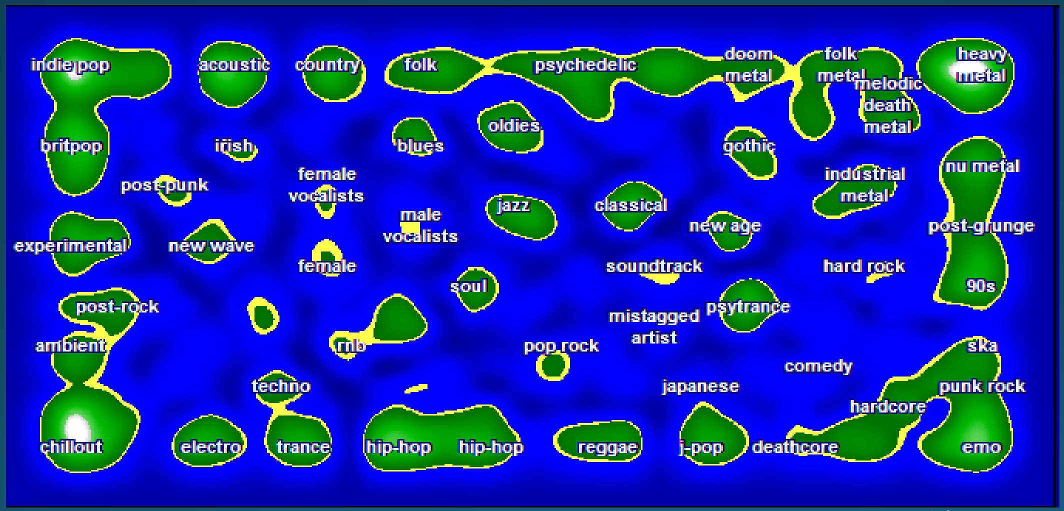
© oSkope, listio

Perceptive dimensions



Maps (3)

2. Data-based dimensionality reduction and visualization
   - Data-dependent but typically global (not query-dependent)
   - Self-organizing feature maps, multidimensional scaling, etc.



© lastFM

   - Information loss inevitable with dimensionality reduction
   - Dimensions usually lack interpretability

Map to known structure

Goal: answers to the query are mapped to the individual components of a known structure

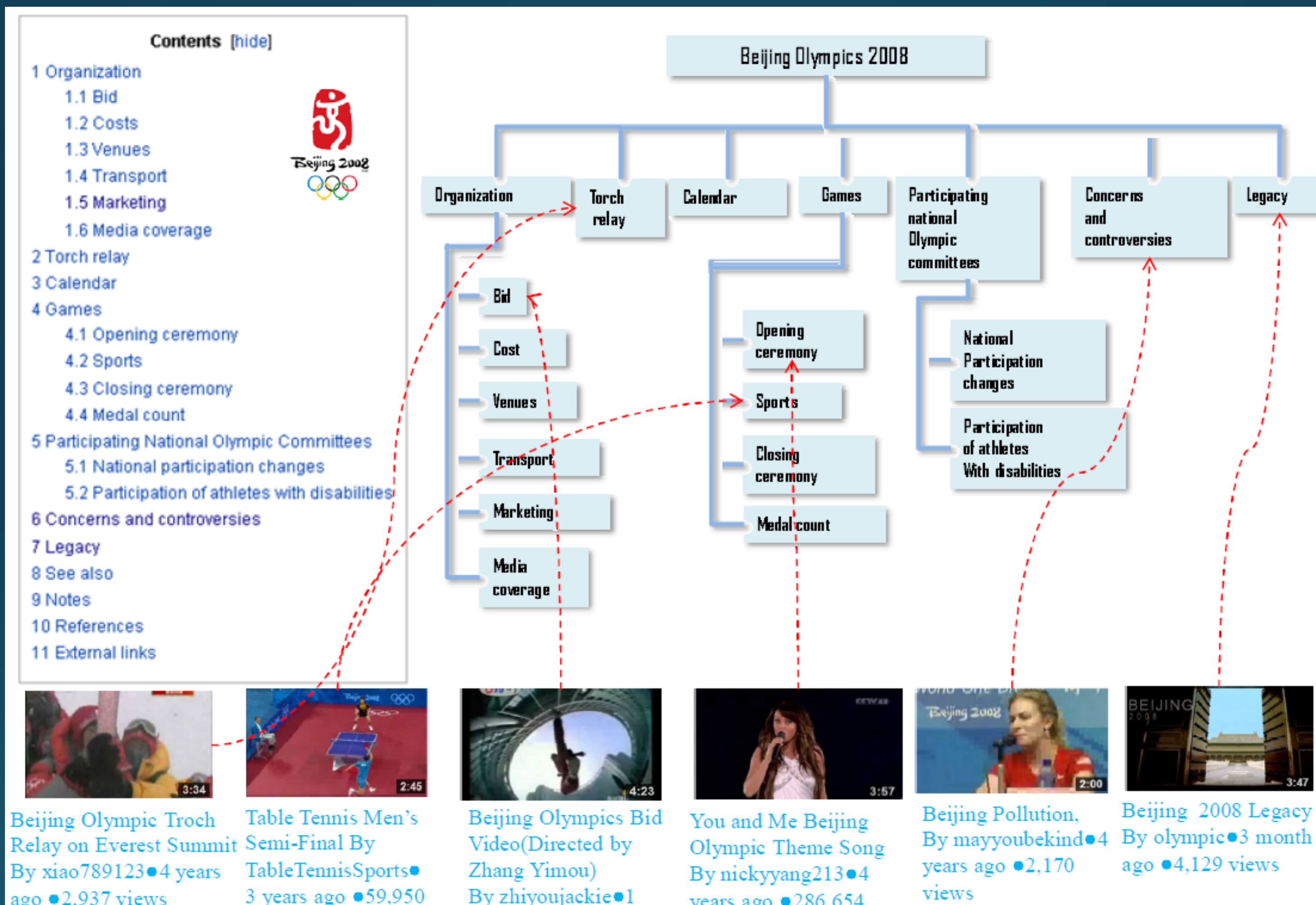
Interesting because allows to 'fill in' the 'slots' of prior, authoritative conceptual structure

- Example
  - Structure the videos received as answers to query for an event according to a predefined hierarchy of topics taken from Wikipedia, see [TJN14]

Principle

- 'Projection' of a content collection on an existing structure
- Prior availability of the known structure is a key assumption

Map to known structure [TJN14]



[TJN14]

Map to known structure [TJN14]

1. Retrieval of videos relevant to the general query
2. Evaluation of relevance video ↔ node, based on
   - Textual similarity: text of the node ↔ text surrounding the video
   - Visual similarity: near-duplicates
3. Specificity video ↔ node: a video should illustrate a node if it is relevant mostly to that node
4. Diversity: a node should not be illustrated with redundant videos (near-duplicates)

Return a knowledge structure

Goal: from the answers to a query, identify a previously unknown structure and map the answers to that structure

Interesting because should return ‘what is there to know about’ the query, for any query and not just predefined ones

Principle

- Tentative: key concept identification, construction of knowledge structure, map answers to the structure
- Possible use of ‘hints’ (e.g. known partial structure)

Potential applications

- Support understanding of (novel) concepts
- Track in time changes in the structure associated to a query

Complex queries

Concept detectors as queries

What is relevant?

- Classically: simple similarity (originating in 'query by keyword')
- Needed: complex relevance measures, learning based

Goal: find documents containing occurrences of a complex concept that cannot be found by simple similarity to query

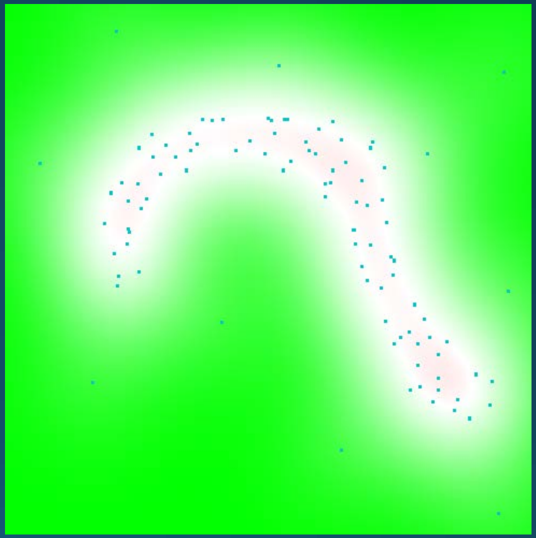
Interesting because allows to employ complex, query-dependent relevance measures

- Examples
  - Find sequences showing taxis picking clients in a video surveillance database
  - Find buildings surrounded by (flammable) scrubland in remote sensing data

Concept detectors as queries (2)

Principle

- Relevant ≡ one side of detection boundary
- Requires training data (possibly enlarged via interactions, e.g. relevance feedback)
- Detector training, query = detector, apply detector to database, return results



Challenges

- Training data for user queries
- Detectors for complex domains e.g. structured, heterogeneous data
- Scalable detection (partly addressed in work on scalable relevance feedback)

Structures as queries

- Goal: answers to the query are 'filled in' structures based on a structure used as query
  - Example: prior conceptual structure for "Beijing Olympic games"
- Should allow for more specific or more generic (or partially-specified) structures
  - Examples
    - "Beijing Olympic games"
    - "Olympic games"
    - "recent international sports competition"
- Challenges
  - Library of structure components supporting query construction
  - Scalability

Conclusion

Search results ranking became a bottleneck in MIR

Going beyond

- Complex answers: provide more comprehensive information
- Complex queries: allow for more sophisticated questions

Major challenges

- Off-line and online data mining for complex answers and queries
- Knowledge mining from multimedia content
- Scalability of complex operations on large multimedia collections

References

[Bro02] Andrei Broder. 2002. A taxonomy of web search. *SIGIR Forum* 36, 2 (September 2002), 3-10. DOI=10.1145/792550.792552

[DB14] Xuan Hong Dang and James Bailey. 2014. Generating multiple alternative clusterings via globally optimal subspaces. *Data Min. Knowl. Discov.* 28, 3 (May 2014), 569-592. DOI=10.1007/s10618-013-0314-1

[HKL12] Alan Hanjalic, Christoph Kofler, and Martha Larson. 2012. Intent and its discontents: the user at the wheel of the online video search engine. In *Proc. 20th ACM intl. conf. on Multimedia* (MM '12). ACM, New York, NY, USA, 1239-1248. DOI=10.1145/2393347.2396424

[IRMMPL14] Bogdan Ionescu, Anca-Livia Radu, María Menéndez, Henning Müller, Adrian Popescu, and Babak Loni. 2014. Div400: a social image retrieval result diversification dataset. In *Proc. 5th ACM Multimedia Systems Conf*. (MMSys '14). ACM, New York, NY, USA, 29-34. DOI=10.1145/2557642.2563670

[LGOZ09] Reinier H. van Leuken, Lluis Garcia, Ximena Olivares, and Roelof van Zwol. 2009. Visual diversification of image search results. In *Proc. 18th intl. conf. on World wide web* (WWW '09). ACM, New York, NY, USA, 341-350. DOI=10.1145/1526709.1526756

[LKM10] Mathias Lux, Christoph Kofler, and Oge Marques. 2010. A classification scheme for user intentions in image search. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '10). ACM, New York, NY, USA, 3913-3918. DOI=10.1145/1753846.1754078

[PC10] Timo Pröscholdt, Michel Crucianu. 2010. On Finding Complementary Clusterings, In *Proc. European Symp. on Artificial Neural Networks* (ESANN 2010), Brügge, Belgium, April 2010, pp. 487-492.

[TCTWDK13] Jaime Teevan, Kevyn Collins-Thompson, Ryen W. White, Susan T. Dumais, and Yubin Kim. 2013. Slow Search: Information Retrieval without Time Constraints. In *Proc. Symp. on Human-Computer Interaction and Information Retrieval* (HCIR '13). ACM, New York, NY, USA, 10 pages. DOI=10.1145/2528394.2528395

[TJN14] Song Tan, Yu-Gang Jiang, and Chong-Wah Ngo. 2014. Placing Videos on a Semantic Hierarchy for Search Result Navigation. *ACM Trans. Multimedia Comput. Commun. Appl.* 10, 4, Article 37 (July 2014), 20 pages. DOI=10.1145/2578394