Search-Based Relevance Association with Auxiliary Contextual Cues

Chun-Che Wu, Kuan-Yu Chu, Yin-Hsi Kuo, Yan-Ying Chen, Wen-Yu Lee, Winston H. Hsu
NTU MiRA, National Taiwan University, Taiwan

2013.10.07
Challenge

Constructing image retrieval models from huge collections, 23M history queries, images and clicks, to measure relevance of any new coming image-query pairs in an online system (< 12 seconds).

**Input:** image & query

suri and katie cruise

**Output:** relevance score

0.68

Database

? How relevant
Strategies and Framework

Input: image-query pair

suri and katie cruise

Output: relevance score 0.68

Ensuring scalability
Leveraging prior knowledge
Strengthening popular queries
Dealing with unseen data

Evaluated by ranking results of a given query:

\[ DCG_{25} = 0.01757 \sum_{i=1}^{25} \frac{2^{rel_i} - 1}{\log_2(i+1)} \]
The query is relevant to the image, if the other visually similar images are associated with similar queries.

1M Visual Words (SIFT)

Visually similar images as candidates

Image-query pair

Search

baby shower centerpieces

baby shower centerpieces

baby shower ideas...

diaper cakes...

drones...

$\sim(t_q, t_i)$
Search-Based Approach

- Retrieving candidates by visual (CBIR) & text (TBIR)
- Similarity are weighted by reliability of candidates (#clicks)

\[ relevance = \sum_{i \in C} sim(v_q, v_i) \cdot sim(t_q, t_i) \cdot click, \]

q: query
C: top ranked images

<table>
<thead>
<tr>
<th>Method</th>
<th>Initial</th>
<th>Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCG@25</td>
<td>0.469</td>
<td>0.484</td>
</tr>
</tbody>
</table>
Speed Up by Indexing

- Why search? Flexible for indexing: inverted-index, KD-tree, ...
- Computation & memory cost will **not surge with incremental data size** -> **Scalable** (< 1 sec/query)

<table>
<thead>
<tr>
<th>Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual word 1 🌟</td>
</tr>
<tr>
<td>Visual word 2 🌟</td>
</tr>
<tr>
<td>Visual word 3 🌟</td>
</tr>
<tr>
<td>Visual word 4 🌟</td>
</tr>
<tr>
<td>...</td>
</tr>
<tr>
<td>Visual word K 🌟</td>
</tr>
</tbody>
</table>

Query image

<table>
<thead>
<tr>
<th>Inverted list</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image ID</td>
</tr>
<tr>
<td>Image ID</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>
Majority in Visual Consistency

Images of the query “Waldo Alabama”

Visually consistent images have higher relevance to the query

Visual Similarity as transition probability (p)

Random Walk
Relevance Propagation by Visual Similarity

\[ s = (\alpha P + (1 - \alpha)v1^T)s, \]

where \( P(i, j) = \frac{\text{sim}(i, j)}{\sum_i \text{sim}(i, j)} \)

\( s \): Score, 
\( v \): Prior knowledge, 
\( P \): Transition probability

15.78% relative improvement

<table>
<thead>
<tr>
<th>Method</th>
<th>Initial</th>
<th>Search</th>
<th>Propagation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCG@25</td>
<td>0.469</td>
<td>0.484</td>
<td>0.543</td>
</tr>
</tbody>
</table>
Popular Queries in Image Search

- 8/15 trending image searches* are **Persons**
- 31.8% Bing dev image queries are **Persons**

**Persons** are very popular search queries

* Bing trending image searches on Sep. 24
Learning Identity Classifiers by Names and Faces

• **Name Detection:**
  – *Collections of celebrity names (2,221)*
  – Mutual combinations of First-Name (1,164) and Last-Name (1,681)

• **Identity Bank:** 6,762 identity models
  – Training by 35,092 face-name pairs
    in **Bing training image set**

• **Challenges**
  – Accuracy
  – Persons out of Identity Bank

**Queries**
- **Tyler Swift** album
- highlights **Michael Jordan**
- **Barack Obama** election

*Cmlab* Communications & Multimedia Lab.

*http://www.posh24.com/celebrities/*
Improve Classification Accuracy

- Identity Bank with background models (IB)

Image-Query Pair

Barack Obama  Michelle Obama

Randomly selected background models

Hypothesis testing statistically significant?

s: classification score by the model of detected name

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>IB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCG@25 (I)</td>
<td>0.352</td>
<td>0.496</td>
</tr>
</tbody>
</table>

\( S' = \frac{S}{r} \)

(I): names included in Identity Bank
Persons out of Identity Bank

- Face number as reference (FN) for smoothing
  - Once a name is detected in a tag, the associated image should comprise at least one face.
  - More names are detected, more faces are expected to appear in image content.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>IB</th>
<th>FN</th>
<th>IB+FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCG@25 (N)</td>
<td>0.481</td>
<td>0.496</td>
<td>0.508</td>
<td>0.516</td>
</tr>
<tr>
<td>DCG@25 (I)</td>
<td>0.352</td>
<td>0.496</td>
<td>0.500</td>
<td>0.510</td>
</tr>
</tbody>
</table>

45.71% relative improvement

(N): partial names appear in name list
(I): names included in Identity Bank
Unseen Names: More Auxiliary Cues

- Detect names out of name list (non-celebrities)

Queries usually contain names if the retrieved results contain social websites

<table>
<thead>
<tr>
<th></th>
<th>Name List</th>
<th>Facebook</th>
<th>LinkedIn</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>0.967</td>
<td>0.73</td>
<td>0.959</td>
<td>0.795</td>
</tr>
<tr>
<td>Recall</td>
<td>0.279</td>
<td>0.585</td>
<td>0.438</td>
<td>0.509</td>
</tr>
</tbody>
</table>

> 20% improvement in recall of name detection
Ongoing and Future Work

• Seeing the unseen data
  – Tag expansion: Snippet, Hash tags
  – Topic modeling

• Describing the unseen data
  – Embedded contexts: Angelina (female) vs. Michael (male)
  – Attributes

• Search or Classification? Query-dependent strategy
  – General queries: search-based
    • Few training data; minor improvement & less scalability by learning
  – Trending queries: classification-based
    • Rich training data; significant improvement & better user experiences by learning
Summary

• We propose an image-query relevance measurement approach considering four major strategies,
  - Ensuring scalability
  - Leveraging prior knowledge
  - Strengthening popular queries
  - Dealing with unseen data

• We demonstrate the efficiency (< 2 seconds/query) and 16% relative improvement in DCG compared to the original ranking results
Acknowledgements

• Thanks for Microsoft Research and Bing for organizing the challenging and great event!

• With the research supports in part by
  – Quanta Research Institute (cloud cluster)
  – Chunghwa Telecom Laboratories
  – Microsoft Research Asia
Thank you for your attention

NTU MiRA, National Taiwan University, Taiwan