Visual Query Suggestion and Diversification of Image Search Results

Helge Holzmann
holzmann@L3S.de
Outline

• Visual Query Suggestion
  • State of the Art / Motivation
  • Approach: Visual Query Suggestion (VQS)
  • Interface
  • Evaluation

• Visual Diversification of Image Search Results
  • State of the Art / Problem
  • Approach: Visual Diversification
  • Clustering Techniques
  • Evaluation

• Combination

• Conclusion

• Discussion
Visual Query Suggestion: Towards Capturing User Intent in Internet Image Search

ZHENG-JUN ZHA, University of Science and Technology of China and National University of Singapore
LINJUN YANG, TAO MEI, MENG WANG, Microsoft Research Asia
ZENGFU WANG, University of Science and Technology of China
TAT-SENG CHUA, National University of Singapore
XIAN-SHENG HUA, Microsoft Research Asia

Query suggestion is an effective approach to bridge the Intention Gap between the users’ search intents and queries. Most existing search engines are able to automatically suggest a list of textual query terms based on users’ current query input, which can be called Textual Query Suggestion. This article proposes a new query suggestion scheme named Visual Query Suggestion (VQS) which is dedicated to image search. VQS provides a more effective query interface to help users to precisely express their search intents by joint text and image suggestions. When a user submits a textual query, VQS first provides a list of suggestions, each containing a keyword and a collection of representative images in a dropdown menu. Once the user selects one of the suggestions, the corresponding keyword will be added to complement the initial query as the new textual query, while the image collection will be used as the visual query to further represent the search intent. VQS then performs image search based on the new textual query using text search techniques, as well as content-based visual retrieval to refine the search results by using the corresponding images as query examples. We compare VQS against three popular image search engines, and show that VQS outperforms these engines in terms of both the quality of image suggestion and the search performance.
Visual Query Suggestion

• State of the Art
  → Visual Queries as list of keywords
  → Hard to express image queries in precise words
Visual Query Suggestion

• State of the Art
Visual Query Suggestion

• State of the Art
  → Intention Gap
Visual Query Suggestion

• State of the Art
  ➔ Textual query suggestion
    ▪ Based on top documents retrieved by initial query / search logs / click-through data
Visual Query Suggestion

- Approach VQS
  - Novel query interface
  - Textual suggestions with representative images
  - Text-based search
  - Visual search result refinement by means of suggested images
Visual Query Suggestion

• Approach VQS
  → Textual suggestions based on tags from Flickr
    ▪ Efficient, without performing search for the initial query
  → Using tagged images as representatives
    ▪ Clustering
Visual Query Suggestion

• Finding Keywords
  → Satisfying relatedness, informativeness:
    ▪ E.g., Apple: Fruit, Computer, Smartphone vs. Smartphone, iPhone

• Finding representative images (exemplars)
  → Affinity Propagation (AP) Clustering

• Refining text-base search results using image
  → Rank by visual similarity to selected image suggestion
    ▪ Characteristics, e.g., color moment, edge distribution, wavelet texture
    ▪ Basic reranking utilizing selected representative
    ▪ One-class SVM based on selected visual suggestion cluster
Visual Query Suggestion

• Interface
Visual Query Suggestion

• Evaluation
  ➔ Quality evaluation: Engine1, Engine2, VQS
    ▪ 30 professional, 10 average users rate systems
    ▪ Much better, better, same, worse, much worse
    ▪ Individual query, overall
    ▪ Result: VQS outperforms existing search engines
  ➔ Performance evaluation: Baseline, TQS, VQS1, VQS2
    ▪ VQS1: basic reranking, VQS2: one-class SVM reranking
    ▪ 20 subjects label relevance level
    ▪ Highly relevant, relevant, irrelevant
    ▪ NDCG@k
Visual Query Suggestion

- Evaluation
Paper 2

WWW 2009 MADRID!

Track: Rich Media / Session: Tagging and Clustering

Visual Diversification of Image Search Results

Reinier H. van Leuken
Universiteit Utrecht
Utrecht, the Netherlands
reinier@cs.uu.nl

Lluis Garcia
Yahoo! Research
Barcelona, Spain
lluis@yahoo-inc.com

Ximena Olivares
Universitat Pompeu Fabra
Barcelona, Spain
ximena.olivares@upf.edu

Roelof van Zwol
Yahoo! Research
Barcelona, Spain
roelof@yahoo-inc.com

ABSTRACT

Due to the reliance on the textual information associated with an image, image search engines on the Web lack the discriminative power to deliver visually diverse search results. The textual descriptions are key to retrieve relevant results for a given user query, but at the same time provide little useful information when making visual comparisons. models deployed on the Web and by these photo sharing sites rely heavily on search paradigms developed within the field Information Retrieval. This way, image retrieval can benefit from years of research experience, and the better this textual metadata captures the content of the image, the better the retrieval performance will be.

It is also commonly acknowledged that a picture has to be
Visual Diversification of Image Search Results

• State of the Art
  → Ranking reflects similarity between query and image’s metadata
  → Topical (textual) diversity / no visual diversity
  → Too many different results, due to too general query (intention gap)
Visual Diversification of Image Search Results

• State of the Art
  →Topical diversification
Visual Diversification of Image Search Results

- Visual Diversification Approach

- Tiger print mammal
- New car model at expo
- Behind bars
- Oldtimer in street
- Black mammal
Visual Diversification of Image Search Results

• Visual Diversification Approach
  → Refine of original search results by means of visual clusters
  → Clustering based on visual similarity
  → 6 visual features:
    ▪ Color histogram
    ▪ Color layout
    ▪ Scalable color
    ▪ CEDD
    ▪ Edge histogram
    ▪ Tamura
Visual Diversification of Image Search Results

• Clustering Techniques
  → Folding
    ▪ Respects original ranking on selecting representative, requires $sd$
    ▪ *Sufficiently dissimilar* ($sd$): mean distance of all images to *avg. image*
    ▪ Clusters formed using nearest neighbor rule
  → MaxMin
    ▪ Aims for maximum visual diversity
    ▪ Each representative has the largest minimum distance to all the others
  → Reciprocal Election
    ▪ Images vote for representative reciprocally ($m$ most similar)
    ▪ Cluster formed around representative $R$ with images having $R$ in top-$m$
Visual Diversification of Image Search Results

• Evaluation
  → 25 textually ambiguous queries
  → 50 textually non-ambiguous queries
  → 8 assessors to cluster top 50 results based on visual characteristics → 200 clusterings (ground truth)

→ Performance bounds (Baseline)
  ▪ Lower bound: interassessor variability
  ▪ Upper bound: Assessor clusters vs. randomly created

→ Two Metrics
  ▪ Fowlkes-Mallows Index
  ▪ Variation of Information
Visual Diversification of Image Search Results

• Evaluation

Reciprocal election achieves the strongest overall performance
Combination

• Query with visual suggestion; diversify image search results
  → Most obvious
  → Not reasonable, as suggestions are already clustered/diversified

• Base VQS suggestions on diversified search results
  → Suggestions with established TQS techniques
  → Cluster results as proposed in Paper 2
  → Use visually diversified cluster representatives as suggestions
  → Perform VQS reranking based on clusters

→ Advantages:
  ▪ Get rid of the dependency on Flickr
  ▪ Suggestions and result diversification combined in one interface
Conclusion

• Visual Query Suggestion
  → Visual query suggestions incorporating Flickr tags
  → Refines initial search result by exploiting visual information
  → VQS outperforms popular image search engines in terms of quality and performance

• Visual Diversification of Image Search Results
  → Compared three clustering techniques for visual diversification
  → Evaluation using two metrics
  → Reciprocal election achieves the strongest overall performance

• Perfectly combinable
  → VQS incorporating visual diversified image search results
Discussion

Thank you for your attention!