PageRank for Product Image Search

Yushi Jing\textsuperscript{1}  Shumeet Baluja\textsuperscript{2}

\textsuperscript{1}College of Computing
Georgia Institute of Technology
Atlanta GA

\textsuperscript{2}Google, Inc.
1600 Amphitheater Parkway
Mountain View CA

WWW 2008, April 2008, Beijing, China
To incorporate visual aspects of images into selecting of images for describing products to buy

Figure: Google Product Search
Outline

1 Background
   - Current Image Search Methods
   - Applying Object Detection to Image Search
   - Detecting and Applying Low-Level Features

2 Approach and Algorithm
   - Reranking Results From Image Search
   - Computing Low-Level Features
   - From Features to Similarity
   - From Similarity to Centrality
   - From Centrality to Reranking

3 Experimental Results
   - User Evaluation (Subjective)
   - Click Study (Objective)
Outline

1. Background
   - Current Image Search Methods
     - Applying Object Detection to Image Search
     - Detecting and Applying Low-Level Features

2. Approach and Algorithm
   - Reranking Results From Image Search
   - Computing Low-Level Features
   - From Features to Similarity
   - From Similarity to Centrality
   - From Centrality to Reranking

3. Experimental Results
   - User Evaluation (Subjective)
   - Click Study (Objective)
Popular search engines for images do not use image features to rank images

Instead, the following are commonly used:

- anchor-text
- image name
- surrounding text on the webpage
Reasons for Text-Based Image Search

- Text-based search is well-studied and successful
- Object recognition still largely unsolved
- Computational complexity of computer vision tasks
Problem With Not Looking At Image Itself

- Results are often inconsistent and uncontrolled in terms of quality and in terms of content
Outline

1 Background
- Current Image Search Methods
- Applying Object Detection to Image Search
- Detecting and Applying Low-Level Features

2 Approach and Algorithm
- Reranking Results From Image Search
- Computing Low-Level Features
- From Features to Similarity
- From Similarity to Centrality
- From Centrality to Reranking

3 Experimental Results
- User Evaluation (Subjective)
- Click Study (Objective)
Previous Method to Use Vision to Rank Images

The algorithm

- Construct models of categories of objects trained from top search results
- Rerank images based on their fit to the model

The problem

- Assumption of one homogeneous object category per query is unrealistic
- Can potentially maximize relevance but not diversity
Moving Away from Object Detection and Towards Local Features

- Authors’ approach does not rely on first detecting *objects*
- Instead, low-level features that are invariant to degradations such as scale and orientation are used
Outline

1 Background
   - Current Image Search Methods
   - Applying Object Detection to Image Search
   - Detecting and Applying Low-Level Features

2 Approach and Algorithm
   - Reranking Results From Image Search
   - Computing Low-Level Features
   - From Features to Similarity
   - From Similarity to Centrality
   - From Centrality to Reranking

3 Experimental Results
   - User Evaluation (Subjective)
   - Click Study (Objective)
Goal: Find Common Features Among Images

Difficulties: Common features may be...

- in any orientation (rotated)
- anywhere in the image
- at any scale (relative size)
- not the main focus of the image (in the background)
- a non-standard color

Figure: Similarity measurement must handle potential rotation, scale and perspective transformations.
Outline

1 Background
- Current Image Search Methods
- Applying Object Detection to Image Search
- Detecting and Applying Low-Level Features

2 Approach and Algorithm
- Reranking Results From Image Search
- Computing Low-Level Features
- From Features to Similarity
- From Similarity to Centrality
- From Centrality to Reranking

3 Experimental Results
- User Evaluation (Subjective)
- Click Study (Objective)
Task from User Perspective

What does the user want?

1. Use existing Google image search algorithm to find the top $k$ images given a verbal query
2. Rerank top $k$ images to maximize relevance and diversity
Goal of Image Search Engines

- Retrieve image results that are relevant
- Retrieve image results that are diverse enough to cover variations of visual or semantic concepts
- Here: Reduce top 1000 to representative 10
Graph Model

Model the *imaginary user behavior* given the visual similarities of the images to be ranked.

- Treat images as web documents
- Treat similarities as probabilistic *visual hyperlinks* (vislinks)
- Estimate the probability of images being visited by a user following these vislinks
- Images with more estimated visits are ranked higher
Vislinks Versus Real Links

- Related web documents are connected by manually defined hyperlinks.
- For images, authors compute vislinks explicitly as a function of visual similarities.
- Idea: a user views one image, other similar images may also be of interest.
- If image $u$ has a vislink to image $v$, then there is some probability that the user will jump from $u$ to $v$.
  - *Random Surfer Model*
Outline

1 Background
- Current Image Search Methods
- Applying Object Detection to Image Search
- Detecting and Applying Low-Level Features

2 Approach and Algorithm
- Reranking Results From Image Search
- Computing Low-Level Features
  - From Features to Similarity
  - From Similarity to Centrality
  - From Centrality to Reranking

3 Experimental Results
- User Evaluation (Subjective)
- Click Study (Objective)
Similarity Measurement

- Global features such as color histograms and shape analysis are too restrictive for the breadth of images that need to be handled
- Local descriptors contain a richer set of image information and are relatively stable under different transformations
- Examples of local features:
  - Harris Corners
  - Scale Invariant Feature Transform (SIFT)
  - Shape Context
  - Spin Images

Figure: Example of use of corners (CSE 152)
Use of Local Features

- Authors use SIFT features with a Difference of Gaussian (DoG) interest point detector and orientation histogram feature representation as image features.

**Figure:** Image taken from Alice Chu and Sparta Cheung from CSE 190a
Difference of Gaussian Interest Point Detector

- DoG interest point detector builds a stack of scaled images by iteratively applying Gaussian filters to the original image.
- Adjacent Gaussian images are subtracted to create DoG images.
Using DoG Image Pyramid

- Characteristic scale associated with each of the interest points can be estimated by finding the local extrema over the scale space.
- Interest points located at local extrema of 2D image space and scale space are selected.
- Gradient map is computed for the region around the interest point and divided into a collection of subregions.
- Orientation histograms can then be computed.
- Final descriptor is a 128 dimensional real valued vector by concatenating 4x4 orientation histogram with 8 bins.
Outline

1 Background
- Current Image Search Methods
- Applying Object Detection to Image Search
- Detecting and Applying Low-Level Features

2 Approach and Algorithm
- Reranking Results From Image Search
- Computing Low-Level Features
- From Features to Similarity
- From Similarity to Centrality
- From Centrality to Reranking

3 Experimental Results
- User Evaluation (Subjective)
- Click Study (Objective)
Defining Similarity

Given two images $u$ and $v$, and their corresponding descriptor vector $D_u = (d_1^u, d_2^u, ..., d_{128}^u)$ and $D_v = (d_1^v, d_2^v, ..., d_{128}^v)$, define similarity between two images as the number of interest points shared between two images divided by their average number of interest points.
Outline

1 Background
   - Current Image Search Methods
   - Applying Object Detection to Image Search
   - Detecting and Applying Low-Level Features

2 Approach and Algorithm
   - Reranking Results From Image Search
   - Computing Low-Level Features
   - From Features to Similarity
   - From Similarity to Centrality
   - From Centrality to Reranking

3 Experimental Results
   - User Evaluation (Subjective)
   - Click Study (Objective)
Abstract Problem Statement

Given a graph with vertices and a set of weighted edges, define and measure the “importance” of each of the vertices

- vertices = images
- edge weights = similarity
Importance of Images

Definition of importance: eigenvector centrality

- Eigenvector Centrality is a method to combine importance of a vertex with those of its neighbors in ranking
- A vertex closer to an important vertex should rank higher than others
Eigenvector Centrality

- Defined as the principal eigenvector of a square stochastic adjacency matrix

Figure: Adjacency matrix for unweighted graph
Eigenvector Centrality (cont.)

- $P$ is a matrix where the probability of moving from image $i$ to image $j$ is $p_{i,j}$
- $P$ is symmetric and stochastic, where each row and column sums to 1
- Matrix constructed from the weights of the edges in the graph
- Ranking scores correspond to probability of arriving in each vertex by traversing through the graph
- Decision to take a particular path defined by weighted edges

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,j} & \cdots \\ p_{2,1} & p_{2,2} & \cdots & p_{2,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{i,1} & p_{i,2} & \cdots & p_{i,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix}$$
Image Rank

Image Rank (R) is defined as the solution to the following equation

\[ R = S^* \times R \]

1. \( S = \{ S_{uv} \} \) where \( S_{uv} \) is visual similarity between image \( u \) and image \( v \)
2. \( S^* \) is the normalized, symmetrical adjacency matrix \( S \)
   - Similarities are assumed to be commutative
   - Repeatedly multiplying \( R \) by \( S^* \) yields the dominant eigenvector of the matrix \( S^* \)

Algorithm to compute \( R \):
   - Guess \( R_0 \)
   - \( R_t = S \times R_{t-1} \)
Convergence

- Image rank converges only when $S^*$ is aperiodic and irreducible
- Incorporating a damping factor $\lambda$ into the original equation, given $n$ images, $R$ is defined as

$$R = \lambda S^* \times R + (1 - \lambda)p$$

follow links
jump anywhere

where $p = |\frac{1}{n}|_{n \times 1}$

This is similar to the random surfer model
Outline

1 Background
- Current Image Search Methods
- Applying Object Detection to Image Search
- Detecting and Applying Low-Level Features

2 Approach and Algorithm
- Reranking Results From Image Search
- Computing Low-Level Features
- From Features to Similarity
- From Similarity to Centrality
- From Centrality to Reranking

3 Experimental Results
- User Evaluation (Subjective)
- Click Study (Objective)
Application Algorithm

- Computationally infeasible to compute similarities for all images indexed by Google
- Instead, precluster web images (based on metadata)
- Rely on existing commercial search engine for initial grouping of semantically similar images
- Given a query, extract top $N$ results returned, create graph of visual similarity on the $N$ images
- Compute image rank only on this subset
Queries with Homogeneous Visual Concepts

- All images look somewhat alike
- Proposed approach improves the relevance of the search results
- Achieved by identifying the vertices that are located at the center of weighted similarity graph

Figure 4: Since all the variations (B, C, D) are based on the original painting (A), A contains more matched local features than others.
Queries with Heterogeneous Visual Concepts

- Example queries are Jaguar and Apple
- The approach is able to identify a relevant and diverse set of images as top ranking results

Figure 2: Many queries like “nemo” contain multiple visual themes.
Example: Monet paintings
Outline

1 Background
   - Current Image Search Methods
   - Applying Object Detection to Image Search
   - Detecting and Applying Low-Level Features

2 Approach and Algorithm
   - Reranking Results From Image Search
   - Computing Low-Level Features
   - From Features to Similarity
   - From Similarity to Centrality
   - From Centrality to Reranking

3 Experimental Results
   - User Evaluation (Subjective)
   - Click Study (Objective)
Image Collection and Queries

- Images collected directly from the web
- Used 2000 most popular product queries (from Google product search)
- Users have strong expectations of the type of results returned
Extraction of Images and Construction of Graphs

- For each query, top 1000 search results extracted
- Similarity matrix constructed by counting number of matched local features for each pair of images
Minimizing Irrelevant Images: User Study

- Designed to study a conservative version of relevancy of the ranking results
- Mixed top 10 selected images using authors’ approach with top 10 images from Google, removing duplicates and presented to user
- Asked the user which images are least relevant to the query
- More than 150 volunteer participants chosen, and 50 queries used

Table: “Irrelevant” images per product query

<table>
<thead>
<tr>
<th></th>
<th>Image Rank</th>
<th>Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among top 10 results</td>
<td>0.47</td>
<td>2.83</td>
</tr>
<tr>
<td>Among top 5 results</td>
<td>0.30</td>
<td>1.31</td>
</tr>
<tr>
<td>Among top 3 results</td>
<td>0.20</td>
<td>0.81</td>
</tr>
</tbody>
</table>
Another View of the Results

**Table: “Irrelevant” images per product query**

<table>
<thead>
<tr>
<th></th>
<th>Image Rank</th>
<th>Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 3 results</td>
<td>0.20</td>
<td>0.81</td>
</tr>
<tr>
<td>Next 2 results</td>
<td>0.10</td>
<td>0.50</td>
</tr>
<tr>
<td>Next 5 results</td>
<td>0.17</td>
<td>1.51</td>
</tr>
</tbody>
</table>
Outline

1 Background
   • Current Image Search Methods
   • Applying Object Detection to Image Search
   • Detecting and Applying Low-Level Features

2 Approach and Algorithm
   • Reranking Results From Image Search
   • Computing Low-Level Features
   • From Features to Similarity
   • From Similarity to Centrality
   • From Centrality to Reranking

3 Experimental Results
   • User Evaluation (Subjective)
   • Click Study (Objective)
Click Study: Setup

- Collected clicks for top 40 images presented by Google search results on 130 common product queries
- For the top-1000 images for each of the 130 queries, rerank them according to the approach described
- To determine if the approach would improve performance, examine the number of clicks each method received from only the top-20 images
Click Study: Results

The images selected by authors’ approach to be in the top-20 results would have received approximately 17.5% more clicks than those in the default ranking.