Multimedia Systems and Applications

Hierarchical Clustering-Based Navigation of Image Search Results

James Wang

Based on ACM Multimedia 2008 paper and poster by:

Ding et al., Hierarchical Clustering-Based Navigation of Image Search Results, in Proceedings of ACM Multimedia 2008, pp. 741 – 744, October 26–31, 2008

Problem and Idea

- **Problems:**
  - *Typically, an image search engine, such as Google or Yahoo, returns a ranked list of search results according to their relevance to a given query.*
  - *However, due to the query's polysemy, the results always contain multiple topics and they are mixed together.*

Figure 1. Top 16 returns of query "apple" in Google's image search engine.

- **Idea**
  - *HiCluster - an effective framework based on a hierarchical clustering structure in order to organize the search results convenient for users browsing.*

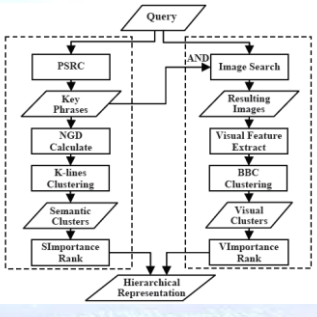
Existing Approaches and Their Problems

- **Traditional Content-Based Image Retrieval (CBIR) uses only visual features to cluster images.**
  - Hard to get consistent clusters from semantic perspective.
- **Deng et al. [1] proposed a hierarchical clustering method using visual, textual and link analysis and finally obtained different semantic clusters.**
  - Could not present a semantic explanation of each cluster, and so the meaning of each cluster is difficult to understand.
- **Igroup [2] used a clustering algorithm for web image search results.**
  - the possibly consistent clusters on semantics are organized with an arbitrary order.
  - there are still many similar key phrases on semantic level.
  - Since the visual feature is not used, the coherence on visual appearance and the representative thumbnails are not available in the application.

HiCluster Approach

- **Query related key phrases are extracted by exploring search results from general textual search engines and K-lines-based clustering algorithm is utilized to cluster the key phrases into semantic clusters.**
- **The resulting images corresponding to each key phrase are grouped into some visually coherent clusters.**
  - Bregman Bubble Clustering (BBC) algorithm is applied to partially group images to discard some scattered noisy ones.
- **A novel UI is designed to make the best of the hierarchical clustering results.**
  - Displays the semantic clusters by putting the appropriate titles together.
  - Presents the visually and semantically coherent clusters under each title by assigning some representative thumbnails.

Hierarchical Clustering Algorithm



Semantic Clustering (Key Phrases Extraction)

- **First, the key phrases are created by the page search results clustering method (PSRC) [3].**
  - We first extracts all possible phrases (n-grams) from the results returned by a textual Web search engine, and then calculates several properties for each phrase such as phrase frequencies, document frequencies, and phrase length and more.
  - A regression learning model from training data is applied to combine these properties into a salience score.
  - The key phrases are ranked by the salience scores, and the top-ranked phrases are taken as key phrases.

Semantic Clustering (K-lines-based clustering)

- **Given the key phrases, the K-lines-based clustering algorithm [4] is selected to organize them into some semantically consistent clusters.**

1. Compute Similarity matrix: $A = (\exp(-NGD(x,y)^2 / 2\delta^2))$ where $\delta^2 = [mean(NGD(x,y))]^2 / 2$
2. Compute Laplacian matrix: $L = S^{-1/2} A S^{-1/2}$ where $S_{ii} = \Sigma_j A_{ij}$ and $S_{ij} = 0$ for $i \neq j$
3. Compute top k eigenvectors $V \in \mathbb{R}^{n \times k}$
4. Cluster V using K-lines[4].

Figure 3. Framework of K-lines-based clustering

- **Normalized Google Distance (NGD) as Eq.1 is used to measure the semantic relevance between any phrase-pair.**

$$NGD(x,y) = \frac{\max\{\log f(x), \log f(y)\} - \log f(x,y)}{\log N - \min\{\log f(x), \log f(y)\}} \quad (1)$$

- **Clusters are ranked in descending order of their semantic importance defined as:**

$$SImportance(m) = \frac{1}{N_m} \sum_{i=1}^{N_m} ResultSize(i) \quad (2)$$

7

Visual Clustering

- **A new clustering scheme, Bregman Bubble Clustering (BBC) [6], is used to visually cluster the images corresponding to each key phrase.**
  - The algorithm obtains the dominant clusters by partially grouping images in the whole set while discarding some scattered noisy ones.
  - From the statistical perspective, it is reasonable to believe that the dominant clusters as main visual representations can reflect the semantics of the given key phrase more efficiently.
  - After the visual clustering is finished, images closest to the cluster center are chosen as its representative images, and all images in the cluster are sorted in descending order of their distance from the cluster center.
- **Different visual clusters are ranked in descending order of their importance defined as:**

$$VImportance(n) = ClusterSize(n) / \sigma_n \quad (3)$$

8

A fast iterative solution for BBC

- **Given k centers from n data points and a desired threshold s of each cluster size, there are three steps:**
  - (1) assigning each point to the nearest center,
  - (2) picking points closest to their centers first until s points are picked, and
  - (3) updating the centers.
- **These three steps are repeated until there is no change in assignment between two successive iterations.**
- **To address BBC's sensitivity to the initial seeds, Pressurization [6] is introduced.**
  - A parameter set of$\{s_j\}$, where $s_j = s + [(n-s) \cdot r^{j-1}]$ is used in jth iteration, r ∈ (0,1) controls the increasing rate of pressure.

9

User Interface

Figure 4. Snapshot of UI for HiCluster

10

Cluster Navigation

Figure 5. Visual clusters navigation UI

11

Experiments

- **UI satisfaction and effort:**

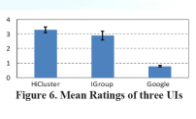


Figure 6. Mean Ratings of three UIs

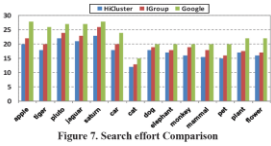


Figure 7. Search effort Comparison

- **Search Performance:**

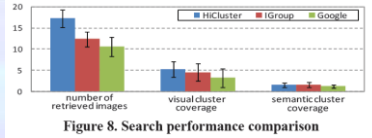


Figure 8. Search performance comparison

12

Conclusion

- **An effective framework for the hierarchical clustering-based navigation of image search results (HiCluster) is proposed.**
  1. the K-lines-based semantic clustering organization, which is performed by putting the semantic similar key phrases together, provides users a concise understanding of the query semantic meanings.
  2. the BBC based visual clustering for the images in each semantic sub-cluster presents users a more informative and convenient navigation by putting the visual similar images together, and tagging them with a representative thumbnail.
  3. a friendly UI is designed to make the best of the hierarchical clustering solution.
- **Comprehensive evaluations are performed to demonstrate the effectiveness of the framework.**