Google Image Swirl

A Large-Scale Content-Based Image Visualization System

Yushi Jing [*] Henry Rowley Jingbin Wang
David Tsai Chuck Rosenberg Michele Covell
Google Inc.
Mountain View, CA, USA

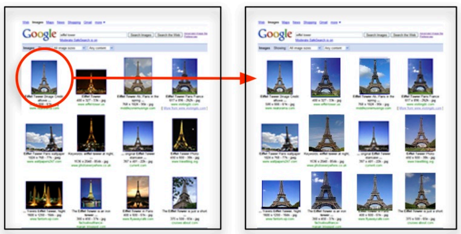

Figure 1: Current query- and content-based Web image search engines present search results as a relevance-ordered list.

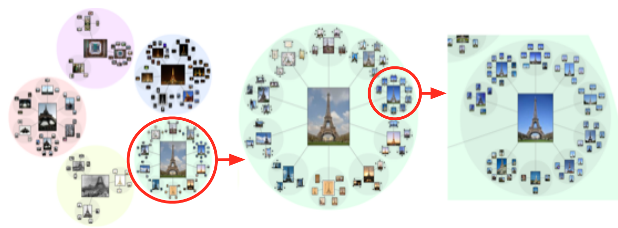

Figure 2: Presenting image search results as an exemplar-tree.

ABSTRACT

Web image retrieval systems, such as Google or Bing image search, present search results as a relevance-ordered list. Although alternative browsing models (e.g. results as clusters or hierarchies) have been proposed in the past, it remains to be seen whether such models can be applied to large-scale image search. This work presents **Google Image Swirl**, a large-scale, publicly available, hierarchical image browsing system by automatically group the search results based on visual and semantic similarity. This paper describes methods used to build such system and shares the findings from 2-years worth of user feedback and usage statistics.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Clustering, Retrieval Models; I.4.10 [**Image Representation**]

1. INTRODUCTION

Current Web image search engines, such as Google or Bing Images, represent search results as a *relevance-ordered list*, as shown in Figure 1. Such list representation is efficient to compute, easy to interpret, and useful when the desired image is among the top search results. However, as text queries are often too simple to capture user intention, and meta-data such as anchor-text are too sparse to describe the images, the desired image (or its near-duplicates) often receives low relevance-scores and are positioned far down in the search results. In such cases, users have no choice but to sequentially browse through a potentially very large number of thumbnails before the desired image is found.

This work studies the feasibility of *automatically* organize the search results as a *exemplar-hierarchy*, based on visual similarity, for large-scale Web image search. An example of our browsing model for the query "Eiffel Tower" is shown in Figure 2. The representations used are organization tools that people are familiar with: examples include the organization of files on a computer and the organization of books in a library. Compared with relevance-ordered list, hierarchy-based browsing model provide users with a global *visual summary* of the content, so users have more information to decide on where to go next given the currently selected document. Also, as a tree can be represented as a planar graph, one can efficiently compute a two dimensional visualization layout.

Although exemplar-hierarchy has been proposed previously as an alternative browsing model to image search [8, 2, 1, 7, 5], due to the limited scale of the proposed systems and related experiments, it remains to be seen whether such methods can be applied to large-scale image search, and whether users would find such system useful. Afterall, due to the semantic gap between the visual representation of the image and the higher level semantics they are associated with, it is not clear one can automatically find obvious and generally agreed-upon dimensions to split the image datasets. Also, due to users' familiarity with the interface of popular Search engines, it is not clear whether typical users would find an alternative interface intuitively and useful.

[*]Please direct all correspondence to jing@google.com. We thank Stephen Holiday's help with the system integration. A live demo can be found at http://cpl.cc.gatech.edu/projects/VisualSynset/.

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

This work presents new evidence on the feasibility of hierarchy-based Web image browsing systems by 1) developing a *hierarchy-based* image browsing system supporting 400K common Web queries and making it publicly available [4], and 2) collecting user feedback and analyzing how the system is being used within its two years life-span.

2. GOOGLE IMAGE SWIRL

Exemplar-hierarchy for 400K Web queries

We collected 400,000 popular queries used on Google image search. For each query, we collected the top 1000 image search results. For each image, we generated various features including color, edge, texture, local features, face signatures and etc. The features are quantitized and L1 Hash is then applied to make the sparse features dense, and KPCA with Histogram Intersection Kernel is used to further reduce dimensionality and place the data in a Euclidean feature space. Pairwise image similarity is computed by applying L2 distance to the features. We then perform clustering to partition the search results into hierarchical clusters, each associated with a representative, or exemplar, image. The hierarchical clusters for each query are pre-computed. For more detail, see [6].

Browsing interface

After hierarchical clustering has been performed, the results of an image search query are organized in the structure of a tree. A number of options exist for how to present such a tree to the user. Beyond the typical layered diagram used to illustrate tree data structures, there are many options in the literature, including using hyperbolic geometry to better utilize space, and a variety of approaches based on tree-maps. In this work, we used balloon-tree layout in which each layer of the tree is arranged radially around its parent. When the user selects a branch of the tree to explore, it is re-scaled to simulate the "zoom-in" effect as shown in Figure 2. The re-arrangement is animated to allow the user to follow the change without getting lost.

3. EXPERIMENTS AND CONCLUSIONS

We selected 1426 of the popular keywords as test queries. We used a combination of Web text and human raters to obtain a set of high-level class labels for images in the search results. We first measure the correlation between visual distance (derived from image features) and semantic distance derived from class labels, with results shown in Figure 3 (right). This result shows there is positive correlation between semantic and visual distance , which is consistent with recent similar experiments using data from ImageNet [3].

Next, we evaluate the quality of clustering by testing the discriminative power of a linear SVM classifier trained on the top-level clusters, similar to those used in ImageNet [3]. The goal is to test whether a given image should belong to the cluster or not. An exemplar-hierarchy is computed for each query. We evaluate the classification results by AUC (the area under the ROC curve). The results are shown in Figure 3 (left). We can see the AUC scores of most clusters are above 0.95, which shows the high visual compactness of the clusters. The hierarchy-based image browsing system was made publicly available from 11/2009 to 19/2011. and usage statistics is represented in Table 1. Users also has the option to rate their browsing experience on a scale of 1 to 5, and our system received an average of 4.6 (5 best).

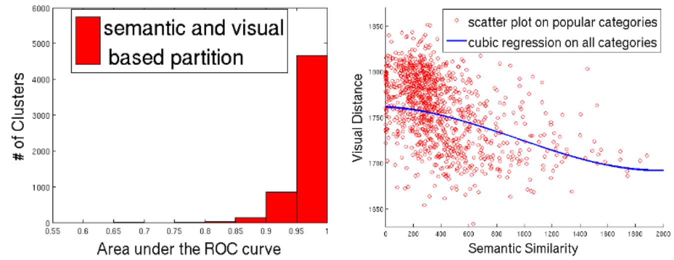

Figure 3: Are under the ROC curve

Usage statistics	mean	std
1) Thumbnails viewed per query	60.92	46.50
2) Interaction per query	4.32	7.84
3) Landing page selections per query	0.62	1.73
4) Session length (seconds)	62.76	288.94

Table 1: Usage statistics

In summary, we developed a large-scale content-based image visualization system based on exemplar-hierarchy. Our experiment results showed positive correlation between semantic and visual distance, and we also received positive feedback from the users.

4. REFERENCES

[1] D. Cai, X. He, Z. Li, W.-Y. Ma, and J.-R. Wen. Hierarchical clustering of WWW image search results using visual, textual and link information. In *Proceedings of the 12th annual ACM international conference on Multimedia*, MULTIMEDIA '04, pages 952–959, New York, NY, USA, 2004. ACM.

[2] J. Chen, C. Bouman, and J. Dalton. Similarity pyramids for browsing and organization of large image databases. In *Proc. of SPIE/IST Conf. on Human Vision and Electronic Imaging III*, 1999.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[4] Y. Jing, H. Rowley, C. Rosenberg, J. Wang, and M. Covell. Google image swirl, a large-scale content-based image browsing engine. In *Demo at IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[5] S. Krishnamachari and M. Abdel-Mottaleb. Image browsing using hierarchical clustering. *Computers and Communications, IEEE Symposium on*, 0:301, 1999.

[6] D. Tsai, Y. Jing, Y. Liu, H. A.Rowley, S. Ioffe, and J. M.Rehg. Large-scale image annotation using visual synset. *ICCV*, 2011.

[7] S. Wang, F. Jing, J. He, Q. Du, and L. Zhang. Igroup: presenting web image search results in semantic clusters. In *Proc. SIGCHI conference on Human factors in computing systems (CHI)*, pages 587–596, New York, NY, USA, 2007. ACM.

[8] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '03, pages 401–408, New York, NY, USA, 2003. ACM.