Metadata Visualization for Image Browsing

Mira Dontcheva
University of Washington
mirad@cs.washington.edu

Maneesh Agrawala
Microsoft Research
maneesh@microsoft.com

Michael Cohen
Microsoft Research
mcohen@microsoft.com

ABSTRACT
We present an interactive visualization approach for browsing a collection of images and their taxonomy. Using mass-spring simulation we generate a layout that clusters images with similar metadata and provide the user with a number of interaction mechanisms for browsing the metadata clusters. It is our hope that this type of interactive visualization will accelerate image search and further aid users in better understanding the contents of a collection of images.

KEYWORDS: Image Browsing, Taxonomy Visualization, Visual Query Refinement

INTRODUCTION
A picture can be worth a thousand words. Unfortunately finding that *perfect* picture to convey those thousand words is often very difficult. Typically the user cannot inspect all images returned for a given query and thus iteratively searches through the database by examining some images and then revising the query. Some image databases today, such as Getty Images [5] or Flickr [4], are organized using a taxonomy of metadata in which every image is associated with one or more string descriptors, or tags. Although tags can help in suggesting ways for narrowing the search, there may be several ways to refine the query and the user may have to cycle through all possible variations to find an appropriate image.

This paper presents an interactive visualization system that makes it easy for the user to browse a collection of tagged images. The system employs a mass-spring simulation to display the images in relationship to all of their tags and provides interaction mechanisms for visually refining queries. It is our belief that displaying the images according to the structure of the metadata can aid the user in finding that *perfect* picture.

Previous work [9] has shown that grouping images according to metadata can be a powerful mechanism for enhancing image browsing. Rodden's work on analyzing image search characteristics [7] shows that viewers prefer topical layout clustering to one based on image similarity. We draw inspiration from their work and focus on generating layouts in which images that share a tag are displayed close together. Treemaps and bubblemaps [1] [3] are well suited for

Figure 1: We create a layout that describes the metadata structure present in 100 images corresponding to the query "universityofwashington". The user highlights three tags: campus, uw, and fog. The images outlined in red are part of both the "campus" and "uw" image clusters, while the image outlined in white is part of all three.

this type of data, however these algorithms require a strict metadata hierarchy. Robertson et. al.'s work [6] on displaying intersecting hierarchies addresses the need for visualizations of information that is part of several intersecting hierarchies, however their work focuses on exploring one entity at a time while we want to visualize many images that are part of many hierarchies simultaneously. Chang and Legget [2] use streaming collage and interaction to provide the user with a comprehensive understanding of a collection of images. As the number of images grows, an interactive approach can quickly become burdensome and an appropriate layout is necessary for efficient browsing. Mass-spring simulations are commonly used for graph layout [8] and we found them appropriate because the relationships between the images and tags form a graph. While the algorithm itself is not novel, our design of the link lengths and their interaction with user feedback to support image browsing is new. We also explore visually refining queries through the manipulation of image clusters.

OVERVIEW
Our system includes two components, a layout algorithm and a set of user interactions. We use flickr.com's [4] image database to test our approach. Flickr is a photograph sharing website that allows users to not only upload and share their images with the public but also to tag their photos with any

number of string descriptors. Flickr provides an API that enables direct access to all publicly available images and their associated tags.

Layout

We initialize the position of each photograph by randomly placing it in a grid of size $\sqrt{n}$ x $\sqrt{n}$, where n is the number of images. We then use a mass-spring simulation to generate a layout to better convey the tag structure within the image collection. In the simulation, all the forces are generated by damped springs. A spring exists between any two images that share a tag and the rest length of that spring is proportional to the number of images corresponding to the tag. We formulate the rest length, L, of a spring between images i and j with tag t as $L(i, j) = k * N(t)$, where k is a constant that controls the density of the layout and $N(t)$ is the number of images associated with t. The quality of the layout depends on the specification of the rest length of each spring. We found that the linear formulation is good for minimizing total area, while still displaying the metadata structure within the image layout. Quadratic formulations, such as $L(i, j) = k * N(t)^2$ create more defined clusters at the cost of sparser layouts, while a square root relationship generates more uniform layouts because the springs are closer in length. Despite appropriate spring lengths, our approach is not successful in clustering all images that share a tag close to one another. Images that are part of many large tag clusters may be pulled to any one of the clusters causing a scattering for any particular set. To generate a layout we evolve the simulation using the Runge-Kutta method with a time step of 0.075. Each image has mass of 100 and the spring and damping constants for all springs are 10 and 100. The simulation converges in under a minute for 200 images on a 2.8 GHz machine with 512MB of RAM.

To show which images are associated with which tags, we display the tags with the images. Each tag is placed in the center of its associated image cluster. We check for overlapping tags and perturb the tags to minimize overlap.

User Interface

To help the user in browsing the collection, we provide mechanisms for highlighting and removing images. Instead of focusing on any individual image, the user interacts with the collection through the tag structure. The user interface consists of a text box for writing queries and two buttons, "Select" and "Prune".

Structure Understanding. We aid the user in discerning which images correspond to a tag through rollovers and highlighting. When the user moves the cursor over a tag, all images that are not part of the tag's image set are dimmed. If the user clicks on the tag, each relevant image is highlighted with a semi-transparent colored overlay specific to the tag. To see which tags correspond to an individual image, the user holds the cursor over the image.

Query Refinement. To facilitate query refinement, we provide two image removal techniques. Once the user has highlighted an image set by clicking on its tag, he can chose to discard it or he can remove all images but the ones highlighted. We provide two buttons: "Select"- for keeping the highlighted images, and "Prune" - for discarding those highlighted. When the user removes images, the simulation cuts those links associated with the discarded images and generates a new layout. Note that we do not change the spring length once the initial layout has been created but rather let the simulation settle to a configuration of lower energy. We use animation to help the user maintain object constancy. Drastic changes in layout are only present when the user selects distant clusters and collapses them together. Usually the clusters maintain their structure as the springs have large damping forces and keep the images in the existing formations.

CONCLUSIONS AND FUTURE WORK

We have presented an interactive visualization system for browsing images and their associated tag structures. Using mass-spring simulation we create a layout that clusters images with similar tags and provide the user with mechanisms for refining queries through pruning tag image sets. It is our hope that this type of interactive visualization will make image search faster and easier. Next, we plan to add the ability to expand tag clusters by adding new images to the layout. We expect that the tag structure will suggest new queries other than those originally envisioned, and the desire to expand the collection to include new tags or more images with a given tag will be common. User studies will also help us better understand whether this layout is intuitive for users and how easily they understand the pruning and expanding mechanisms. We would also like to carefully analyze the relayout problem and whether users prefer to keep the layout static throughout the pruning and expanding iterations. One challenge for this type of approach is scale. A simulation may not be appropriate for collections much larger than a few hundred images as it may take too long for the layout to stabilize. One approach would be to do rough initial clustering of the images and then apply the mass-spring model to create locally good layouts.

REFERENCES

1. Ben Bederson. Photomesa: A zoomable image browser using quantum treemaps and bubblemaps. *ACM Symposium on User Interface Software and Technology, CHI Letters*, 3(2):71–80, 2001.
2. Michelle Chang, John J. Leggett, Richard Furuta, Andruid Kerne, J. Patrick Williams, Samuel A. Burns, and Randolph G. Bias. Collection understanding through streaming collage. In *Information Visualization Interfaces for Retrieval and Analysis at the Joint Conference on Digital Libraries*, 2003.
3. Gouthami Chintalapani, Catherine Plaisant, and Ben Shneiderman. Extending the utility of treemaps with flexible hierarchy. In *Proceedings of Eighth International Conference on Information Visualization*, pages 335–344, 2004.
4. Flickr. http://www.flickr.com.
5. Getty Images. http://www.gettyimages.com.
6. George Robertson, Kim Cameron, Mary Czerwinski, and Daniel Robbins. Polyarchy visualization: Visualizing multiple intersecting hierarchies. In *Proceedings of SIGCHI*, 2002.
7. Kerry Rodden, Wojciech Basalaj, David Sinclair, and Kenneth Wood. Does organisation by similarity assist image browsing? 2000.
8. Ioannis Tollis, Giuseppe Di Battista, Peter Eades, and Roberto Tamassia. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, 1998.
9. Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. Faceted metadata for image search and browsing. In *Proceedings of SIGCHI*, 2003.