

A Keyword-driven User Interface for Hierarchical Image Browser CAT

Ai Gomi, Reiko Miyazaki, Takayuki Itoh*
Graduate School of Humanities and
Sciences, Ochanomizu University

Jia Li†
Department of Statistics, The
Pennsylvania State University

ABSTRACT

Previously we have presented CAT (Clustered Album Thumbnail), a technique for browsing large image collections, and its interface for controlling the level of details (LOD). CAT applies tree-structured clustering to images based on their keywords and pixel values, and selects representative images for each cluster. A hierarchical data visualization technique displays the tree structured organization of images using nested rectangular regions. Interlocked to the zooming operation, CAT selectively shows representative images while zooming out, or individual images while zooming in.

This poster presents a keyword-driven user interface for CAT. When a user specifies one or multiple keywords on the keyword selection screen, CAT extracts a subset of the tree containing clusters annotated by the user-specified keywords. This interface contributes to filter undesired clusters at an early step of user operation, and makes users easier to focus on their interested images.

1 CAT: A HIERARCHICAL IMAGE BROWSER

Technique for browsing larger collection of images is an interesting topic, and there have been novel works. Here, we suppose that it is not always necessary to show all the images at the beginning of browsing, if we would like to browse thousands of images. We therefore focus on hierarchical image browsers, which first show the representative image of each image cluster, and then users can manually explore each cluster to see each image in the clusters.

We have presented CAT (Clustered Album Thumbnails) [1], a technique for browsing clustered images, and its interface for controlling the level of detail (LOD). CAT first constructs a two-level hierarchy of images as a preprocessing; it first divides images according to their keywords, and then divides again according to contents (colors and textures). CAT then selects representative images for each cluster of images. Finally, CAT displays the hierarchy by using our own hierarchical data visualization technique [2] which represents the hierarchy as nested rectangular regions.

Also, CAT provides a zooming user interface so that users can intuitively focus on images in their interested clusters. While a user zooms out, CAT displays representative images of high-level clusters. Zooming in, CAT displays independent images in each cluster. Figure 1 shows an example of zoom in and out states of CAT.

2 IMPLEMENTATION OF CAT

2.1 Keyword-based Image Clustering

As the first step of preprocessing, CAT constructs clusters of images based on their keywords. Let the whole vocabulary of keywords be V , and the set of keywords for image X_i be W_i , where

$$W_i = \{w_{i,1}, \dots, w_{i,m_i}\}, w_{i,j} \in V \quad (1)$$

and m_i denotes the number of keywords for image X_i . If W_i and W_j are entirely equal, CAT put the images X_i and X_j into the same cluster.

*e-mail: { gomiaai, reiko, itot } @itolab.is.ocha.ac.jp

†e-mail: jiali@psu.edu

Here, displayed sizes of representative images of clusters are approximately proportional to the numbers of images in the clusters. In the other words, it may be inconvenient that representative images are displayed very small when the clusters are too small. To solve the problem, CAT may merge small clusters during the clustering process. Here, CAT calculates distances between all possible pairs of keywords using a natural language processing tool, so that it can select semantically close clusters to merge.

2.2 Content-based Image Clustering

As the second step of preprocessing, CAT further divides images in the clusters generated by the keyword-based clustering, using pixel information. CAT calculates feature vectors based on color and texture. We simply calculate feature vectors from colors and textures of images, and applies bottom-up linkage clustering to divide the images according to cosine between the feature vectors.

2.3 Representative Image Selection

As the final step of preprocessing, CAT selects representative images for each cluster. There is a variety of ways to select representative images, but current our implementation simply select representative images according to pixel information.

In many cases, the image that is closest to the center of the cluster in the feature vector space looks average in the cluster, and therefore the image is preferable as the representative of the cluster. Our implementation therefore simply selects the image closest to the center of the cluster as the representative image.

2.4 Display of Hierarchically Clustered Images

CAT applies our hierarchical data visualization technique [2]. It places a set of images onto a display space based on a bottom-up packing algorithm consists of the following three phases:

Phase 1: CAT first places a set of image thumbnails in a lower-level cluster in grid layout, and encloses them by a rectangular border. It repeats this process for all the lower-level clusters.

Phase 2: CAT then packs and encloses all the rectangles corresponding to the lower-level clusters that belong to the same higher-level cluster by a rectangular border. It repeats this process for each of the higher-level clusters.

Phase 3: CAT finally packs the rectangles of all the higher-level clusters, and encloses them by a rectangular border.

Since CAT places representative images of clusters into the rectangular borders, aspect ratios of the rectangular areas should be close to the aspect ratios of the representative images. For this requirement, CAT calculates the horizontal and vertical numbers of images in the grid layout so that the ratio of the numbers is as close as possible to the aspect ratio of the representative image of the cluster. Also, we arrange the condition of rectangle placement in Phases 2 and 3, so that the aspect ratios of rectangular regions get enough close to the aspect ratios of representative images.

2.5 User Interface for Image Browsing and LOD Control

CAT switches the displaying image according to wheel operation. While zooming out, it displays representative images of higher-level clusters. Zooming in, it switches to representative images of

lower-level clusters, and finally to each image thumbnails. The representative images are displayed inside the rectangular borders of the clusters. CAT stretches the representative images if the aspect ratios of the rectangular borders are not equal to those of representative images.

If the initial viewing configuration zooms out, CAT first loads only representative images from the hard disk drive into the main memory, and then loads each image thumbnails in the focused clusters on the fly, or frees memory space for image thumbnails in defocused clusters. This mechanism is effective for frame rate and memory usage.

3 KEYWORD-DRIVEN USER INTERFACE

After using CAT, we found several problems as follows. First, users may not be interested in all clusters, especially when an image collection includes variety of keywords. In this case, it may be better to filter the clusters so that CAT only display clusters interested by the user. Second, it is often possible that semantically close clusters are distantly displayed. For example, our sample image collection has clusters of "flower", "plant", and "flower and plant". However, it is often possible that the three clusters are distantly placed. When most preferable images of flowers distribute into the three clusters, it may be difficult for users to find the all preferable images.

To solve the problem, we provide a keyword-driven user interface. Figure 2 shows an example of the user interface, which displays a list of keywords annotated onto at least one cluster, used as the initial screen of CAT. Current our implementation selects a representative image for each keyword, and displays onto the display space with the keywords. The implementation also provides a click operation, used for the selection of one or more keywords.

Following is the processing flow of the keyword-driven user interface. After the preprocessing (hierarchy construction and representative image selection), CAT loads the entire hierarchical structure and paths of image files in order: however, it does not load any images themselves at that time.

Let the set of keywords for cluster C_i be W_i , where

$$W_i = \{w_{i,1}, \dots, w_{i,m_i}\}, w_{i,j} \in V \quad (2)$$

and m_i denotes the number of keywords for cluster C_i . Also, let the set of user-specified keywords be S , where

$$S = \{s_1, \dots, s_M\}, s_i \in V \quad (3)$$

and M denotes the number of user-specified keywords.

When a user selects the keywords, CAT constructs a subset tree structure consists of images which have all the user-specified keywords. If the set of keywords W_i includes all keywords in S , cluster C_i will remain in the subset tree; otherwise, C_i will not remain.

This implementation certainly reduces the number of clusters displayed while zooming out, and therefore it makes easier to focus on interested images.

ACKNOWLEDGEMENTS

This work has been partially supported by Japan Society of the Promotion of Science under Grant-in-Aid for Scientific Research (C) No. 18500074.

REFERENCES

- [1] Gomi A., Itoh T., Li J., CAT: A Hierarchical Image Browser Using a Rectangle Packing Technique, *IEEE Symposium on Information Visualization, Poster Session*, 2007.
- [2] Itoh T., Takakura H., Sawada A., Koyamada K., Hierarchical Visualization of Network Intrusion Detection Data in the IP Address Space, *IEEE Computer Graphics and Applications*, 26, 2, 40-47, 2006.

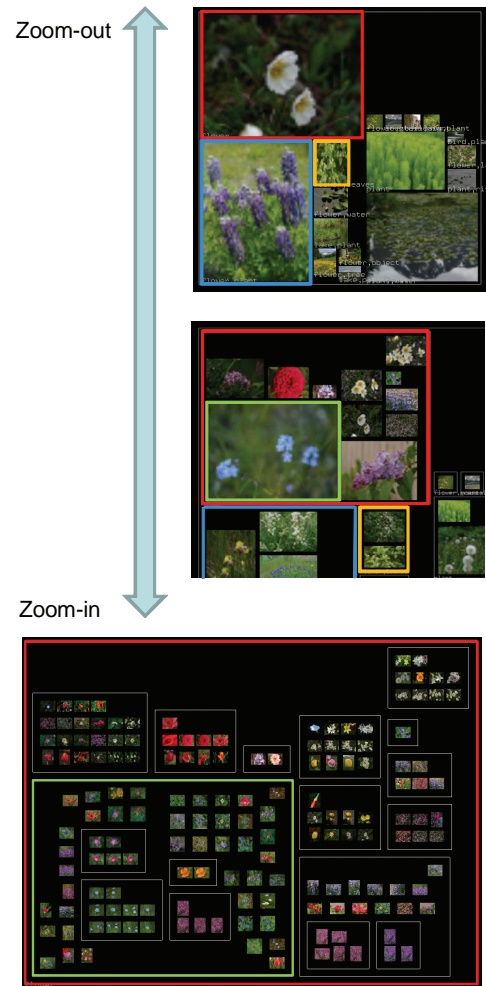


Figure 1: Overview of zooming user interface of CAT. (Upper) While a user zooms out, CAT displays representative images of higher-level clusters. (Center) While a user zooms in, CAT displays representative images of lower-level clusters. (Lower) While a user further zooms in, CAT displays independent thumbnail images.

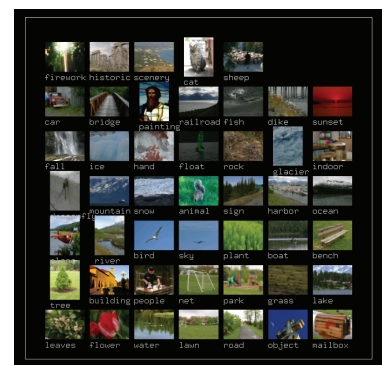


Figure 2: Initial screen for keyword selection.