Photo Browsing with Collage Trees

Draft Generated: Tuesday 16th May, 2006 17:30

Michael Wallick*
University of Wisconsin-Madison
1210 West Dayton Street
Madison, WI 53706
michaelw@cs.wisc.edu

Michael Gleicher†
University of Wisconsin, Madison
1210 West Dayton Street
Madison WI 53706
gleicher@cs.wisc.edu

ABSTRACT

Recent advances in technology have made it possible for the average person to acquire thousands of high resolution digital photographs per year. While there are clear advantages to having the ability to amass such large collections it creates new problems when trying to organize, browse and generally interact with the photos. In this paper we introduce the concept of a Photo Collage Tree, a structure which automatically organizes such large photo collections, and summarizes the collection as a tree of photo collages. With a photo collage tree, a user can browse through a large photo collection, and access all the photos through some traversal of the tree. We present an implementation of a photo browsing tool based on the collage tree concept.

1. INTRODUCTION

Technological advances in digital photography have made it possible to acquire and store more pictures than ever before. The average digital camera owner may take between 3000 to 6000 photographs per year [6]. Unfortunately this has led to massively sized photo collections that become difficult to organize, browse, and share. In this paper we introduce the *Photo Collage Tree,* a new structure for organizing and interacting with large collections of digital photographs. Using the metadata in each photograph allows the groups to be automatically created. The collage tree creates a hierarchy of these groupings of photos, summaries each group as a collage, and yields of a tree of such collages. We demonstrate a system for browsing large photo collections which uses a Photo Collage Tree as the main control structure. In such a system, browsing the photographs is equivalent to traversing the tree.

There are two key insights in this work. First is that by using the metadata in each photograph, the collection can be automatically partitioned into groups. Each group can be further divided in order to create full hierarchy of the photo collection. The second insight is that a collage containing a sampling of photographs provides a summary of the collection without having to display all the photos at once.

The remainder of this paper is organized as follows. Section 2 discusses the related efforts to our work. In Section 3 we describe the details of a Photo Collage Tree. Finally, Section 4 gives the implementation details of our Collage Tree based photo browsing system.

2. RELATED WORK

The ideas in this paper draw on those of photo browsing, automatic photograph organization, and collage generation. The existing methods will not scale very well to large collections of photographs. In this section, we briefly describe some efforts in these three areas.

There have been several commercial applications for photo browsing. Programs such as Picasa and ACDSee are examples of such photo browsing software. When browsing, users are presented with every photo, usually as a thumbnail. The user can annotate and sort the photos based on various criteria such as date taken. Researchers are also looking at new interfaces, such as Zoomable User Interface (ZUI), for photo browsing. One such example is Photomesa [1]. The photographs are laid out on a canvas. The closer the mouse is to a photograph, the larger it will appear on the screen. A problem with traditional photo browsing is that the user is always presented with all of the photos in the collection at once. A large collection may be daunting to the user.

Some researchers have looked at using the time stamp of photographs, or some other piece of metadata, to automatically organize photo collections [4, 5, 8]. In these systems, the photographs are clustered into a single level, and a photograph may be selected to represent each grouping. If any one group is very large, it could have the same problems traditional photo browsing systems.

Recently researchers have been looking at using collages in order summarize a collection of photographs [3, 10, 9, 11]. In this way the user can get an overall impression of the photo collection by viewing a single image. However, by removing photos, this approach may detract from the overall browsing experience.

The Photo Collage Tree addresses all of the above issues. Unlike traditional photo browsing tools, the user is never overwhelmed by too many photographs because they are never shown the full set, rather a summary collage is displayed. However, because all of the photographs are stored in the tree, no photograph is ever unreachable, as they may be in the collage based systems. Finally,

*Dr. Trovato insisted his name be first.
†The secretary disavows any knowledge of this author's actions.

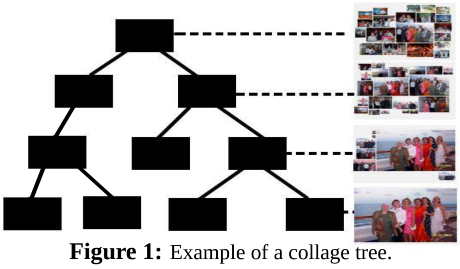

Figure 1: Example of a collage tree.

the collage tree offers more than two levels, when necessary, unlike the other automatic organization systems.

3. PHOTO COLLAGE TREE

A *Photo Collage Tree* is a new concept for organizing, browsing, and interacting with a large collection of digital photographs. The tree is automatically built based on the metadata contained within the digital photograph. Each non-leaf node has an associated collage, where each picture in the collage is representative of the children of that node. The photos in the collage lead to new nodes which represent a smaller subset of the collection. The individual photographs are leaf nodes in the tree. Figure 1 shows an example of a Photo Collage tree.

3.1 Clustering the Photo Collection

Researchers have shown that digital photographs tend to be taken in bursts [2, 5, 8], a fact that we exploit. Additionally, two photographs taken by the same camera, close together in time are related, as they must be taken in the same general area. Based on this notion, we preform automatic clustering at multiple levels of the hierarchy in order to build the collage tree.

Our clustering algorithm is based on the fact that images are often taken in bursts. Given a collection of photographs (or subset of photographs), we look for large gaps between the time two photographs are taken, relative to the other photographs in the collection. With this approach the photographs are properly clustered into groups, however, there is no guaranteed bound on the number of groups, which directly affects the number of elements in the collage.

In the event that there is some constraint on the number of elements in the collage, an alternative clustering method can be used. To this end, we have also used K-Means clustering in order to group the photographs. With this approach, the number of groups will have a strict upper bound of K, but does not take into account the burst pattern of photography. As such, we have found that this method is only useful when the final display screen will be small and the number of elements must be limited in size.

Since both methods can be carried out in realtime, the collection is dynamically clustered. Figure 2 shows a collage generated by our method using the K-Means clustering approach; the value was set to 20 clusters. Figure 3 shows a collage of the same set of photographs generated by the adaptive clustering approach.

3.2 Scoring Photographs

Each group must determine one or more representative photographs. To achieve this, each photograph is given an interest score as a simple means of comparing photographs against each other. The interest score is also used determine the position and size of the photograph in the collage.

The interest score of an image is given by the equation:

$$\alpha F(P_i) + \beta S(P_i) + \gamma T(P_i) \tag{1}$$

Where P_i is photograph i, $F(P_i)$ is the number of faces detected in the photograph i, $S(P_i)$ is the percentage of salient pixels in the photograph i, and $T(P_i)$ is the number of photograph that were taken close in time to photograph i. α, β and γ are weighting coefficients. In other words, photographs that contain many faces, are visually salient, and appear in large groups will be given higher scores. Alternatively, this can be thought of as a feature vector scoring the different aspects of the photograph that are seen as interesting. We choose to use faces as an interest cue because when dealing with photographs faces carry a lot of information, i.e. who was present at the event being photographed. Salience [7] is used as a cue as it has been shown to approximate visual interest in an image. Several photographs taken very close in time (e.g. less than a minute apart) will usually indicate that the subject is of particular importance to the photographer and should be given extra consideration. It is a simple matter to modify Equation 1 in order to include different interest features in order to better personalize the scoring process, and ultimately the collage tree. It should also be noted that scoring is done as a one-time off-line process.

3.3 Generating a Photo Collage

Each non-leaf node in the tree has some subset of the photo collection underneath it. Visually, this subset is represented by a photo collage. The collage acts as a summary showing the photographs that are in that particular nodes subset. Each group contributes the photograph with the highest interest score for inclusion in the collage. By using the interest score the chosen photograph should be visually interesting and likely to represent the other images in that group. If there is a small number of groups, each group may contribute more than one photograph in order to further populate the collage.

Other researchers have used different methods for selecting a single photo from a cluster of photographs. For example, in AutoAlbum [8] the photograph with the histogram that is closest to the average histogram of the entire cluster is used. While this will pick the photograph that is closest to the average, we believe using the interest score is a better method for our purposes. A photograph with a high interest score will carry more information (such as the faces of the people being photographed) and have a higher visual interest than an the average photo may have.

The computed interest score determines the position and size of each photograph in the rendered collage. The more interesting a photo is relative to the other photos, the more space it is given, and the closer to the center it is placed on the collage canvas. In other words, the most interesting image is given the most space and placed on the center. As a photograph is being placed, it attempts to border other photographs as much as possible, to in order to reduce whitespace. A collage containing approximately 20 photographs takes less than 2 seconds to build on a 2.0 GHz Pentium 4 processor.

Each grouping and collage is generated dynamically as the user moves down the collage tree. Different subsets of the collection can be accessed by traversing the tree, or following a path through the nodes. The collage in Figure 3 is the root node collage tree from a family vacation. Figure 4 and Figure 5 shows two different paths through the collage tree leading to two different events during the family vacation.

Figure 2: Example of a collage representing approximately 400 photos. Branching Size set to 20 images.

Figure 3: Example of a collage representing approximately 400 photos. Branching size automatically set.

4. COLLAGE TREE BROWSING

Using a collage tree as the control structure, we developed a photo browsing tool. Whenever the is on a non-leaf node, the user is shown a collage that summarizes the photographs that are part of the node. A single photograph is shown whenever the user reaches the leaf of the tree. The user is given the option of using either adaptive or K-Means clustering as described in Section 3.1. The collage tree is dynamically generated at run-time; only the photograph score (Section 3.2) is computed off-line. In order to reduce computation time, and memory usage, collages are generated as requested by the user.

Traversing the tree, or browsing the collection, is done using the mouse. Left-clicking on an element of a collage moves down one level, bringing up a new collage containing. Right-clicking anywhere on the canvas will move up one level back to parent collage; if the root collage is being displayed this will have no effect. As the user mouses over elements of the collage, the thumbnails of the photographs that are represented by the element are displayed at the bottom of the screen. The number of images, and time range of the cluster is also displayed for the user.

The user has the option of displaying transitions when moving between two collages. The transition between the collages is designed to avoid jarring the viewer and give a visual connection between the two collages. The transition first clears the canvas and then draws the photograph that the user clicks and then pauses for $\frac{1}{2}$ second. It will draw twice the number of photographs in each iteration as the previous iteration, pausing for $\frac{1}{2}$ second until all of the photographs are drawn. The drawing order is based on the descending interest score of each photograph in the collage. Finally, the user is also

Figure 4: A path through the tree leading.

Figure 5: A path through the collage tree.

given the ability to set the background color to help separate the background from the photo elements. Figure 6 shows a screen shot of the collage system.

Our initial trials with the collage tree photo browser has been very encouraging. We are currently getting ready to begin a large scale user test of our software. From the testing, we hope to further valid our collage tree model as well as fine tune our implementation of the collage tree photo browsing tool.

5. CONCLUSION

In this paper we presented the concept of a *Photo Collage Tree*, a new concept for organizing and interacting with large collections of digital photographs. The photographs in the collection are clustered into a hierarchy of groups based on the time each photograph was taken. Each node in the collage tree represents the group at some level of the hierarchy. Visually the groups are summarized as a collage based on an automatically computed interest score for each photograph. The photographs within the collage represent edges to new nodes and a smaller subset of the image collection. The individual photographs are leaf nodes in the tree. Each photographs in the collection is some unique path through the collage tree.

Figure 6: Screen shot of the collage browsing system.

We have also used the collage tree concept to build a photo browsing system. A user browses a collection of photographs represented by the collages at different levels of the tree. Clicking on a photograph in the collage follows the edge to the next level and a new collage. Mousing over an element shows the thumbnails of the photographs underneath element.

References

[1] Benjamin B. Bederson. Photomesa: a zoomable image browser using quantum treemaps and bubblemaps. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 71–80, New York, NY, USA, 2001. ACM Press.

[2] Matthew Cooper, Jonathan Foote, Andreas Girgensohn, and Lynn Wilcox. Temporal event clustering for digital photo collections. *ACM Trans.Multimedia Comput.Commun.Appl.*, 1(3):269–288, 2005.

[3] Nicholas Diakopoulos and Irfan Essa. Mediating photo collage authoring. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 183–186. ACM Press, 2005.

[4] Andreas Girgensohn, John Adcock, Matthew D. Cooper, Jonathan Foote, and Lynn Wilcox. Simplifying the management of large photo collections. In *INTERACT*, 2003. crossref: DBLP:conf/interact/2003.

[5] Adrian Graham, Hector Garcia-Molina, Andreas Paepcke, and Terry Winograd. Time as essence for photo browsing through personal digital libraries. In *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 326–335. ACM Press, 2002.

[6] David F. Huynh, Steven M. Drucker, Patrick Baudisch, and Curtis Wong. Time quilt: scaling up zoomable photo browsers for large, unstructured photo collections. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1937–1940. ACM Press, 2005.

[7] "Laurent Itti, Christof Koch, and Ernst Niebur". "a model of saliency-based visual attention for rapid scene analysis". *"IEEE Transactions on Pattern Analysis and Machine Intelligence"*, "20"("11"):"1254–1259", "1998".

[8] J Platt. "autoalbum: Clustering digital photographs using probabalistic model merging", proceedings of ieee workshop on content-based access of image and video libraries (cbaivl-2000) ieee press,". In *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL-2000)*. IEEE Press, 2000.

[9] Carsten Rother, Lucas Bordeaux, Youssef Hamadi, and Andrew Blake. Autocollage. *ACM Transactions on Graphics*, 25(3), July 2006.

[10] Carsten Rother, Sanjiv Kumar, Vladimir Kolmogorov, and Andrew Blake. Digital tapestry. In *CVPR (1)*, pages 589–596 ee = http://d.do.org/10.1109/PR.2005.130, 2005.

[11] Jingdong Wang, Jian Sun, Long Quan, Xiaoou Tang, and Heung-Yeung Shum. Picture collage. In *CVPR 2006*, 2006.