ICMR 2011

# PhotoCube: Effective and Efficient Multi-Dimensional Browsing of Personal Photo Collections

Grímur Tómasson School of Computer Science Reykjavik University, Iceland grimur09@ru.is

Björn Þór Jónsson School of Computer Science Reykjavik University, Iceland bjorn@ru.is

# ABSTRACT

It has never been so easy to take pictures, and personal image collections have never been so large. Unfortunately, most current photo browsers provide very limited support for effectively navigating image collections. This demonstration proposal describes PhotoCube, a personal image browser based on a multi-dimensional data model similar to the model used in OLAP applications. With PhotoCube, users can tag pictures, structure tags into various hierarchies, and browse images according to any possible perspective. We also describe three demonstration scenarios that show the power, flexibility and scalability of PhotoCube.

# 1. INTRODUCTION

Digital cameras are used all over the planet to capture pictures. Most of these pictures are then stored on personal computers in personal photo collections. Such collections quickly become quite large, as people commonly take a few thousand pictures every year. And for the most enthusiastic, collections can grow orders of magnitude larger.

Current photo browsers provide some support for managing large collections of personal photos. Most systems offer nice facilities to browse photos according to the time line. They also include ways to define folder hierarchies where photos can be assigned; these folders are then surrogates for classifying pictures according to events, people, places. Browsers also offer some support for tagging, either manual or (semi-)automatic (e.g., face recognition). Tagging allows users to later search for the photos containing a certain tag. Finally, browsers often provide very valuable tools for creating slideshows, web-pages or high-quality printed albums. Products like iPhoto and Picasa are good examples.

Such browsers, however, fail to allow users to browse their picture in a flexible manner. While multiple tags can be assigned to a single image, tags are essentially simple attributes, as there is no mechanism for defining any relations between tags. In order to find all the photos where at least one of their friends appears, e.g., users must type in all their names and then search; alternatively, users could think about assigning the tag 'friend' to pictures in addition

Copyright is held by the author/owner(s). *ICMR* '11, April 17-20, Trento, Italy ACM 978-1-4503-0336-1/11/04.

Hlynur Sigurþórsson School of Computer Science Reykjavik University, Iceland hlynur10@ru.is

Laurent Amsaleg IRISA–CNRS Rennes, France laurent.amsaleg@irisa.fr

to their names—all in all, this is quite painful. There is limited or no support for filtering photos according to a value range, for example a time period or brightness. There is also limited support for grouping photos by contents. The commercial tools we have today are thus less than satisfactory.

Some research projects have considered browsing, including PhotoMesa [2], PhotoFinder [3] and Scenique [1]. Overall, these projects are either quite rigid in defining hierarchies and/or tags, or it is quite unclear whether they can handle collections of images of realistic sizes.

In this demonstration proposal, we first describe a multidimensional data model allowing for effective browsing [5]. We then describe the PhotoCube prototype, where a GUI allows users to specify and then visualize pictures according to their will. Finally, we propose three demonstration scenarios that we will share with conference attendees.

## 2. MULTI-DIMENSIONAL DATA MODEL

The multi-dimensional model of data, commonly used in OLAP (On-Line Analytical Processing) applications, has been used very effectively in Business Intelligence applications to view, group and aggregate numerical data, such as sales and profits. While the multi-dimensional model is not immediately suited to photo browsing, we believe that the problems are similar enough to use the concepts and operations of that model as the foundation of a powerful and flexible solution to media browsing [5].

# 2.1 The OLAP Model

The two key concepts of the OLAP multi-dimensional model are numeric *facts* (also called measures) which are categorized by *dimensions*. OLAP users can view facts along any combination of dimensions and, to facilitate analysis of data, the model allows for structuring dimensions in a hierarchical manner. Hierarchies are defined with aggregation operations allowing users to view facts at different granularity levels, and to traverse between levels, e.g., viewing sales of a product or product group per day, week, quarter, or year.

Aggregating data is the key to the main operations OLAP systems provide: slicing and dicing, drilling down, rolling up, and pivoting. Overall, these operations navigate through the existing hierarchies, displaying a particular subset of the data, changing the viewpoint or the level of detail.

### 2.2 From Facts to Photos

Viewing images instead of numeric facts is slightly more complicated (see [5] for more details). Pictures correspond roughly to OLAP facts, but tags (any meta-data) can be associated with those pictures; one photo may be associated with many tags while a particular tag may be associated with many photos. Tags are grouped into tag-sets, which match the OLAP concept of dimensions; these dimensions can further be organized into hierarchies. While some might be predefined (e.g., through taxonomies), tags, tags-sets, and hierarchies can all be arbitrarily defined by users. A simple example would be to create a tag-set with the names of people, and then define two hierarchies of tags organizing, on one hand, family members into a family tree hierarchy, and on the other hand, friends from various places into a friends hierarchy.

OLAP operations are transformed into adding or removing selection predicates applied to the dimensions of a photo collection. Predicates act as filters on tags, tags-sets or hierarchies, and restrict the set of pictures to display. A photo browsing session thus consists of applying filters and retrieving the images that pass through all the applied filters. A drill down operation would, e.g., switch from displaying the pictures according to the continents where they were taken to displaying them according to countries. Note that predicates can do range filtering and filter according to any set of tags or according to any vertex in the existing hierarchies.

In order to facilitate adding automated tagging functionality, the model supports plug-ins through a well defined interface. A plug-in is simply a piece of software that performs some specific analysis of a picture and returns tags, e.g., face recognition and EXIF meta-data extraction.

### 3. THE PHOTOCUBE PROTOTYPE

The PhotoCube prototype has two parts, clearly separated using an API. First, it includes a graphical user interface, which grants users the access to the powerful and flexible browsing operations of the data model, such as the drilldown and roll-up operations, pivoting, and general filtering, via hierarchies and tag-sets. The PhotoCube browser is built using Python and Panda3D, allowing for three-dimensional browsing of the set of pictures of interest. Second, it includes a picture server translating the current browsing state defined through the GUI into complex query to the database where meta-data associated with pictures is stored. Note that PhotoCube is data store independent, but so far the data store showing the best performance is the MonetDB column-store [4], which will be used in the demonstration [5].

Figure 1 shows a sample screen-shot from the PhotoCube browser, with a browsing scenario involving a family photo album. Three dimensions are visible in the figure. On the "up" axis, the family hierarchy is considered with a filter on children, while in the "right" axis, the family hierarchy is considered with a filter on grandparents. Finally, on the "left" axis, a location hierarchy is considered, at the level of places in Iceland. The figure thus shows those photos that have both a child and a grandparent in them, and classifies those photos by the child, grandparent, and place involved. Note that isolating this set of photos, and presenting the set in this manner, is very difficult in traditional browsers.



Figure 1: PhotoCube Scenario (see text for details)

#### 4. DEMONSTRATION SCENARIOS

We propose the following three demonstration scenarios for ICMR attendees. The first photo collection will be a set of about 1,500 photos from a well-known nature trail in Iceland. These will be tagged with information about participants, geography, and other information. We will show how easy it is to navigate through this collection to tell stories, to focus on events, on people, or on places. We will use this collection to discuss the fundamental differences between our data model and previous photo browsers.

The second photo collection will be created dynamically during the ICMR conference. We propose to take many pictures during the conference and to insert these into the collection. These can then be tagged according to either predefined tag-sets and hierarchies, e.g., relating to sessions or papers, or by adding tags and augmenting hierarchies. This will demonstrate the consequences of adding tags, or augmenting hierarchies, on the browsing experience.

The third collection will be used to demonstrate the ability of PhotoCube to cope with scale. We will show its performance when manipulating over 30,000 images. Tags will be randomly generated according to specific distributions in order to stress particular aspects of the underlying prototype.

#### 5. **REFERENCES**

- I. Bartolini. A multi-faceted browsing interface for digital photo collections. In *Proc. CBMI*, 2009.
- [2] B. B. Bederson. PhotoMesa: A zoomable image browser using quantum treemaps and bubblemaps. In *Proc. UIST*, 2001.
- [3] H. Kang and B. Shneiderman. Visualization methods for personal photo collections: Browsing and searching in the PhotoFinder. In *Proc. ICME*, 2000.
- [4] S. Manegold, P. A. Boncz, and M. L. Kersten. Optimizing database architecture for the new bottleneck: Memory access. *VLDB Journal*, 9(3), 2000.
- [5] G. Tómasson. ObjectCube A generic multi-dimensional model for media browsing. Master's thesis, Reykjavik University, Iceland, 2011.