MMM 2015

A Multi-Dimensional Data Model for Personal Photo Browsing

Björn Þór Jónsson[1], Grímur Tómasson[1], Hlynur Sigurþórsson[1]
Áslaug Eiríksdóttir[1], Laurent Amsaleg[2], and Marta Kristín Lárusdóttir[1]

[1] CRESS, School of Computer Science, Reykjavík University, Iceland, bjorn@ru.is
[2] IRISA-CNRS, Rennes, France, laurent.amsaleg@irisa.fr

Abstract. Digital photo collections—personal, professional, or social—have been growing ever larger, leaving users overwhelmed. It is therefore increasingly important to provide effective browsing tools for photo collections. Learning from the resounding success of multi-dimensional analysis (MDA) in the business intelligence community for On-Line Analytical Processing (OLAP) applications, we propose a multi-dimensional model for media browsing, called M^3, that combines MDA concepts with concepts from faceted browsing. We present the data model and describe preliminary evaluations, made using server and client prototypes, which indicate that users find the model useful and easy to use.

Keywords: Photo Browsing; Multi-Dimensional Analysis; Data Model; Graphical User Interface; Evaluation.

1 Introduction

With the recent technological changes, photo collections have been growing very rapidly and there appears to be no end to this growth. This calls for very effective tools for not only *finding* content in those collections, but also gaining *insights* into the collections and *analyzing* them. Search and browsing tools are ubiquitous and to some extent they do help with finding photos in collections, although it can be argued that with the growth in collections the effectiveness of current search and browsing tools is likely to diminish. They offer no support for insight and analysis, however, so what is clearly needed is a new approach to browsing collections. Such an approach has many applications in diverse domains, including professional photo management, online photo stores, digital heritage, and personal photo browsing.

1.1 Background

For traditional databases, the multi-dimensional analysis (MDA) model used in on-line analytical processing (OLAP) was the key to allowing analysis and understanding of large data collections. The MDA model introduced two key concepts that revolutionized users' perception of data, namely *dimensions*, including hierarchies, used for specifying interesting sets of data and *facts*, or numerical attributes, which are aggregated for an easy-to-understand view of the data of interest. These simple concepts put the focus squarely on the value of data items and the relationships that exist between data items.

Based on the resounding success of OLAP applications, it is not surprising that multimedia researchers have studied the application of OLAP to multimedia retrieval for some time (e.g., see [1–3]). The fact that the MDA model is geared towards simple numerical attributes, however, is a serious limitation when it comes to multimedia collections where tags and annotations are a very important part of the meta-data.

The use of tags has been studied in faceted search, however (e.g., see [4–6]). Faceted search uses a single tag-set, but proposes to build multiple hierarchies (or even DAGs) over that tag-set, one for each aspect that could be browsed. These hierarchies are then traversed to interactively narrow the result set, until the user is satisfied. Item counts or sample queries are typically used to present the result while it is very large; when it is sufficiently small it is presented in a linear fashion. A major drawback of the faceted approach, compared to MDA, is the use of a single tag-set; although the hierarchies do help users somewhat to disambiguate the different uses of an ambiguous tag, it is more logical to categorize the tags into different tag-sets.

1.2 Contributions

The major contribution of this paper is a proposal for a new model for photo browsing, which builds on and combines aspects of MDA and faceted search to get the best of both worlds. We call this model the Multi-dimensional Media Model, or M^3; we choose to pronounce this as *emm-cube*, which also refers to the fact that the data model essentially constructs hyper-cubes of photos.

In the M^3 model, photos correspond roughly to OLAP facts, but meta-data items (we call these *tags*, but they can refer to any meta-data, including numerical data, dates and time-stamps, annotations and textual tags) can be associated with the photos; one photo may be associated with many tags while a particular tag may be associated with many photos. Tags are grouped into multiple *concepts*; each concept encapsulates a particular conceptual group of tags, such as people, objects, animals, creation dates, focal length, and so on. Borrowing from faceted search, however, *concepts* are then further organized by building (multiple) browsing *dimensions*, typically structured as hierarchies or DAGs, to facilitate browsing.

The M^3 model thus allows us to define the concepts, according to which photos can be grouped, and the organization of the tags in the concepts. OLAP operations are transformed into adding or removing selection predicates applied to the dimensions of a media collection. Predicates act as filters on tags, concepts or dimensions, and restrict the set of photos to display. A photo browsing session thus consists of: repeatedly adding or removing filters; retrieving the photos that pass through all the applied filters; and displaying the photos dynamically according to the organization of the currently visible browsing dimensions.

2 The M^3 Data Model

This section develops the M^3 model and contrasts it with the MDA model used to view and analyze numerical data.

2.1 Photo Description

The following components of the M^3 model apply to photos and their content.

Object. An object is any entity that a user is interested in storing information about, in this case a photo. Objects correspond to *facts* in MDA, as both represent information users are interested in analysing and both have associated meta-data that further describe them. Unlike facts, however, objects such as photos represent complex content, making aggregation difficult.

Tag. A tag is any meta-data that can be associated with objects. There is no limitation on how many objects a tag can be associated with, nor is there a limitation on how many tags can be associated with a single object. Tags correspond roughly to *members* in MDA.

Location. The location refers to where a particular tag applies to a particular object. The details of the implementation of location may depend on the context, but in the case of photos a bounding box is typically sufficient. A tag that applies to the entire photo, however, does not require a bounding box. There is no counterpart in MDA, as the data items in MDA have no structure.

2.2 Multi-Dimensional Aspects

In this section we define four abstractions that concern categorization and grouping of data, and are at the heart of the multi-dimensional nature of the M^3 model.

Concept. A concept is a set of tags that the user perceives to be related; a conceptual group of tags. In a concept named 'People' the tags can, e.g., be names of people or names of subcategories, say, 'Children' or 'Class Mates'. Some concepts may be entirely user generated, e.g., an 'Event' concept. Others may be based purely on the meta-data associated with each object, e.g., a 'Creation Date' concept. Yet others may partially use automated content analysis, e.g., face identification for the 'People' concept. Any implementation of the M^3 model must therefore support automated media analysis methods.

Concepts are mathematical sets in the sense that tags are distinct. The order of tags, however, can be relevant; consider for example a concept containing creation date tags. Concepts are very similar to *dimensions* in MDA. In both cases the user perceives the tags as being strongly related to each other, and in both cases hierarchies can be built to organize their contents.

Dimension. A dimension adds structure and order to a subset of the tags of a concept. A dimension is typically derived from a single concept and only contains tags from that concept, but may include tags from different concepts.

Each concept has a "default" dimension which includes all the tags in the concept, but may additionally have zero, one, or more associated dimensions. The 'People' concept, e.g., could have one dimension called 'Friends' and another called 'Family', which would typically be largely disjoint but might share some tags. While concepts may or may not be ordered, the vertices of a dimension are explicitly ordered.

The dimension concept is highly similar to the hierarchy concept from MDA, although the dimension concept is more flexible. Hierarchies in the MDA model can represent either level or value based hierarchies, while a dimension may be a more freely structured hierarchy or even a DAG. In both cases, however, a node is the aggregation of its children. In MDA, aggregation is typically based on mathematical functions, such as sum or average, but in M^3 aggregation will typically take some form of grouping.

A node in a dimension may optionally have a title that applies to its children, called a *child category title*; the children are then instances of the category that the title names. To use a standard example, 'Month' could be a child category title, while the months themselves would be the children. A minor difference between the two models is thus that in MDA a column name supplies a level name in a hierarchy, but in M^3 the child category name is applied to a hierarchy node and only applies to the children of that node.

Hypercube. A hypercube is created by selecting and storing information about one or more dimensions which the user wishes to browse objects by.

In MDA implementations, the hypercube is typically stored in a specialized data structure for efficiency. This data structure stores both base facts and pre-calculated aggregations of facts. Due to the differences between objects and facts, however, it is possible that only base data can be stored for the M^3 model and the hypercube will therefore only be conceptual. The representation of complex M^3 hypercybes is a topic for future research.

Cell. A cell in the hypercube is the intersection of a single tag from each of the dimensions in a hypercube. A cell can contain zero or more photos, unlike the MDA model which simply aggregates all the facts corresponding to each cell.

2.3 Retrieval

So far, we have defined abstractions to describe and organize media objects, but now we turn to the retrieval of the objects. Unlike typical search applications, the retrieval is based on browsing, where the retrieved object sets are defined incrementally, based on the user's interest. In the M^3 model, different *filters* can be applied to the various dimensions, resulting in a *browsing state*. We now define these in more detail.

Filter. A filter is a constraint describing a sub-set of photos that the user wishes to see. Each filter applies to a single dimension, but many filters may be applied to the same dimension. All objects that are associated with tags that satisfy the constraint of a filter are said to pass through the filter.

We have defined three different filter variants:

– *Tag Filter:* The tag filter is a filter that selects a single tag from a concept. It is used to retrieve only photos associated with that particular tag.
– *Range Filter:* A range filter applies to ordered concepts and defines a value range by two boundary values, where both boundary values are included in the range. The boundary values themselves need not exist as tags in the concept the filter is applied to.
– *Dimension Filter:* A dimension filter selects a single node in a dimension. The entire sub-structure, or sub-dimension, of that node is said to pass through the dimension filter. A dimension filter on the root of a dimension thus returns the entire dimension, but excludes the tags from the underlying concept that are not part of the dimension.

Note that a filter can be applied to any dimension, regardless of whether that dimension is used for organizing photos in the user interface. Furthermore, a filter continues to restrict retrieval until it is explicitly revoked. A browsing session thus consists of repeatedly applying and/or removing filters and retrieving objects that pass through all applied filters.

The filter concept corresponds to two MDA abstractions. First, it serves the same purpose as *selection*, as both can be used to define a filter to restrict the data retrieved. Second, the filter also corresponds to *page dimension*, as both can be used to restrict retrieval using a dimension not in the cube.

Browsing State. As mentioned above, a browsing session consists of applying or removing filters and retrieving the objects that pass through all the applied filters. The browsing state therefore contains information about filters, tags or sub-dimensions that pass through the filters, and objects associated with these. Note that tags and sub-dimensions are included even when there are no corresponding objects, as the existence of such tags and sub-dimensions is highly useful information for the user. Objects are only returned, however, if they pass through *all* the applied filters. It is sufficient that one tag that the object is associated with passes through each filter, and it is possible that different tags may pass through different filters.

Informally, we can think of each filter as selecting a sub-set of the objects. The objects in the browsing state are then selected by the intersection of the sets passing through the filters. Concepts with no filters are excluded from consideration; if no filters are in effect the browsing state therefore includes all objects. The browsing state corresponds loosely to a *sub-cube* of the hypercube in the MDA model, as the browsing state contains enough data to build a sub-cube.

3 Evaluation

In this section, we discuss our evaluations so far. After briefly describing our prototypes and photo collection, we present a detailed browsing scenario followed by the results of two different user evaluations. Finally, we address one pressing issue in photo browsing, namely that of effective tag generation.

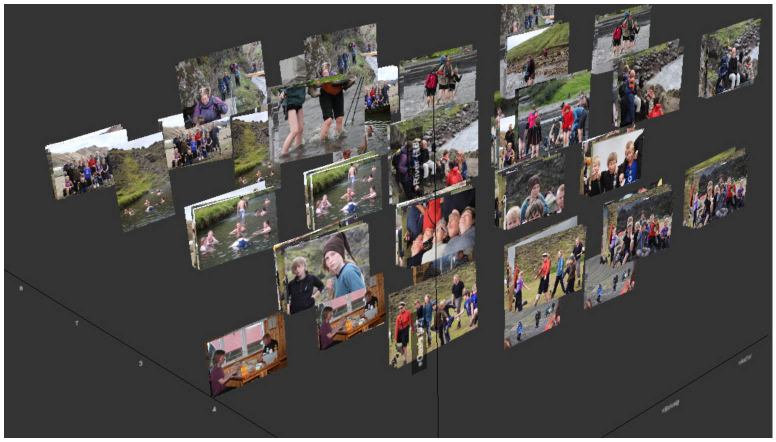

Fig. 1. A three-dimensional browsing state from the scenario.

3.1 Prototypes and Photo Collection

We have developed a media server—called O^3 since it serves hypercubes of objects [7]. It implements most of the M^3 data model, but dimensions are currently restricted to hierarchies over a single concept. A *plug-in* architecture is used to generate tags from Exif metadata and extract faces from photos [8]. In [7], an extensive performance evaluation shows that the server scales well for large photo collections. We have also developed a photo browser—called P^3 since it displays hyper-cubes of photos [9]. P^3 includes a graphical user interface which grants users the access to the powerful and flexible browsing operations of the M^3data model, such as the drill-down and roll-up operations of dimensions, dimension pivoting, and general filtering. Various aspects of these prototypes have been demonstrated at the ICMR, CBMI and MMM conferences [10–12].

The photo collection contained photos from a five day trip from 2010 along a well-known hiking trail. The hiking group consisted of 9 adults (including one of the authors) and 9 children from 5 different families. The collection consisted of $1,140$ images. Aside from 19 meta-data concepts extracted from photo headers, there were 126 tags belonging to 7 concepts with a total of 8 dimensions, in addition to the default dimensions. The concepts were: ‘Events’; ‘Days’; ‘Locations’; ‘People’; ‘Objects’; ‘Animals’; and ‘Impression’ (containing the tag ‘Beautiful’).

3.2 Detailed Browsing Scenario

An adult user is sitting down with her children to recall a hiking trip. She first selects the family dimension (a hierarchy on top of the people concept) as a starting point, and drills down to her own family, which has four members. Then she selects the location dimension as the “up” axis, which has such nodes as “cabin” and “river”. Being a photo nerd, she becomes interested in the light conditions, and selects the aperture value to the third axis. The current browsing state, shown in Figure 1, then has three dimensions, where each cell has one particular family member in one particular location type with one aperture value. Note that photos containing all four family members will show up in four cells (and, if a cabin were situated next to a river, it could show up in 8 stacks); this is an important feature of the model as the photos belong *logically* in all these cells. Then she decides to focus on photos taken at rivers, as she wants to talk about wading the glacial rivers on the hike. As there was one such river every day of the hike, she replaces the aperture value axis with a day axis. She might then go on to consider animals or objects, rotate the cube for a better view, etc.

This scenario demonstrates several common operations in the M^3 model, namely filtering, drilling down into dimensions, and pivoting browsing dimensions. It also shows how well the model uses the screen to indicate why each photo is included in the result. Finally, the scenario illustrates the suitability for storytelling and discovery, which we have experienced vividly in demonstrations.

3.3 Evaluation I: Advanced Users

The first evaluation was a pilot study with experienced computer users, in order to quickly gain insights into the pros and cons of the M^3 data model. For this study we used the same image collection as before. The subjects were five male advanced students of computing. Their experience of image browsers was very varied, but none of the participants had any knowledge of the collection. Aside from questionnaires and an open interview at the end, the focus was on solving specific tasks.

Task Performance: Each participants was asked to perform the following tasks:

1. Show images of kids by location.
2. Show images that contain a sheep.
3. Show images containing hiking shoes which have Aperture value 4 - 5.
4. Show images of people playing football.
5. Show images containing the participating author, grouped by F-number, ISO-Speed and location.
6. Show some images which the participant found to be “cool”.

The purpose of the final task was to allow the users to “play around” with the prototype on their own. The performance on the tasks was noted by the experimenter, using the following performance indicators A (finished without problems), B (finished after experimentation), C (finished with minor help) and F (did not finish).

Table 1 shows the outcome of the task performance evaluation. Each row contains the task performance of one individual participant. The table shows that all participants, except for participant 3, experienced some difficulties with the first two tasks. This is most likely due to the learning curve of the model and prototype. Two users experienced significant difficulties with the first tasks and had to ask for clarifications. The instructions they received, however, consisted solely of reminding them of functionalities of the model, available operations in the browsers and information about the dataset. No direct help instructions were given; yet we do not see any F labels.

	Tasks					
Participant	**1**	**2**	**3**	**4**	**5**	**6**
1	*B*	*B*	*B*	*A*	*A*	*A*
2	*B*	*B*	*A*	*A*	*A*	*A*
3	*A*	*A*	*B*	*A*	*A*	*A*
4	*C*	*B*	*A*	*A*	*A*	*A*
5	*C*	*C*	*A*	*B*	*A*	*A*

Table 1. Eval I: Task Performance.

User Experience Factor	Score	Range
Simple	4.8	3–7
Pliable	5.0	4–6
Easy to use	5.2	4–6
Comfortable	5.2	5–6
Encouraging	5.4	4–6
Enjoyable	6.2	6–7
Imaginative	6.4	6–7
Useful	6.4	6–7
Fascinating	6.6	6–7

Table 2. Eval I: User Experience (7 = best).

Overall, participants experienced the most difficulties with Task 1, as only one user was able to finish this task without problems. In Task 2 the performance was slightly better, but still only one user was able to finish the task without any problems. In Task 3 the performance improved further, as two users were able to finish the task without problems and the others without any assistance. By Task 5 all participants were able to complete the task without any problems, even though that is a relatively difficult task. These results indicated that there was a learning curve for the browser, but that a few browsing sessions might be sufficent to overcome it.

User Experience Factors: Table 2 shows the results of a user experience questionnaire, both the average score and the range of scores. The user experience factors are ordered from the lowest score to the highest score. Overall, we observe that the scores are rather high, ranging from 4.8 to 6.6, but with notable difference between the score for different factors. The factors in the upper half, Simple, Pliable, Easy to use, Comfortable and Encouraging all have a comparably low average score, ranging from 4.8 to 5.6, with a wide range of scores. The widest range is for the Simple factor, where the average value is the lowest and the scores range from 3 to 7. On the other hand, the factors in the lower half, Enjoyable, Imaginative, Useful and Fascinating, all have comparably high average scores of 6.2 through 6.6. Furthermore, the participants all seem to agree on these values, as the range of scores is quite narrow. These results indicate that our participants find the prototype rather complicated and cumbersome to use, while finding it at the same time highly enjoyable, imaginative, and useful.

3.4 Evaluation II: Novice Users

The second evaluation was a more detailed user evaluation with the actual people from the family hike. This is actually a unique evaluation, in the sense that all participants felt as if they were browsing their private image collection. Nine users participated in all, five adults and four children/teenagers, with similar balance between genders. All the participants can be considered novice users.

The experimental protocol was similar as before, except that no predefined tasks were given. Instead, the participants were asked to browse the photo collection freely and make observations about the photos—to "show" the photo collection to the experimenter. This phase lasted 40 to 60 minutes, ending when the participant appeared tired and/or ready to quit. Due to the small size of the user group, we did not compare directly to other browsing tools, but instead asked participants to compare their experience to their regular photo browser.

AttrakDiff Results: The AttrakDiff questionnaire, a standard questionnaire for evaluating software (http://attrakdiff.de/index-en.html), was filled in by seven participants (young children were excluded). The results are interpreted along four different axes: Pragmatic Quality (PQ); Hedonic Quality - Stimulation (HQ-S); Hedonic Quality - Identity (HQ-I); and Attractiveness (ATT). The report states that the PQ and HQ-S values are average, and hence improvements are needed in terms of usability and stimulation. The HQ-I and ATT values are above average, however, meaning that the users found the prototypes attractive and they identified with it. It is possible that these results are biased, however: The fact that users were indeed browsing photos from their own experience, can positively impact the HQ-I value, and the fact that users knew one of the researchers from this trip may have positively impacted the ATT value.

Browsing Questionnaire: We also prepared a questionnaire specifically for photo browsing. The questionnaire contained six questions about the user experience, and the participants responded using a 7-point Lickert scale (1 = I agree; 7 = I disagree). The results are shown in Table 3. Overall, the participants were very positive towards the prototype. In particular, the last question of Table 3 shows that the participants enjoyed using P^3 more than using their regular photo browser. Two participants mentioned that they felt obliged to rate the software as it was at the time of testing, but expressed their expectation that the software would improve with further development and thus would probably deserve an even higher rating in the future. Similar reservations are likely made by other users as well, as the last two questions, which put the prototypes into a more general context, have less positive replies.

Result Summary: It was encouraging to see how quickly this varied set of novice users was able to start using the prototype effectively. An open interview at the end also pointed towards some improvements that could be made in the interface. The measured results from this experiments were not very conclusive, however, but we nevertheless feel that this experiment is worth reporting on as it points to a general methodology for studying personal photo browsing: Find groups with shared experiences and subject them to collections from those experiences.

3.5 Facilitating Tagging

To effectively support media browsing, the M^3 model relies on the existence of tags; when tags exist they can be used very effectively for browsing. In the past, systems relying on tagging have had problems [13] but we believe that several features of our prototypes can help make it very easy for people to *create* the tags, through automated tagging and through user-interface techniques.

Questions About $\mathbf{P}^3$	Average	StdDev
The program helped me recall the hike	1.25	0.46
I enjoyed browsing the photos in this program	1.22	0.67
The program made it easy to browse photos	1.25	0.46
The program made it easy to find photos	1.12	0.35
I would like to show others my photos using this program	1.62	1.19
I enjoyed using this program more than the one I normally use	1.56	1.13

Table 3. Eval II: User Experience (1 = I agree; 7 = I disagree).

This is work in progress, but we already made initial steps towards facilitating tagging. We have, e.g., designed an automated plug-in architecture for tag generation which could be used to implement any sort of analysis technique. During import of photos, tags can be selected that apply to all photos in the import. Furthermore, the new photos are assigned two random numbers that can be used to scatter them on screen, to give a fresh light-board style overview.

Finally, we have designed a drag-and-drop tagging process, which is a new tagging paradigm. For example, a photo, a part of a photo, or a set of photos, may be dragged to any tag that is visible on the screen and dropped there, resulting in a connection between that tag and the photo.

4 Related Work

Many research projects have considered photo browsing and proposed many interesting methods. While the literature is far too vast to be all cited here, we describe some of the most interesting techniques.

PhotoFinder [14] provides boolean search capabilities for finding photos based on several attributes. PhotoMesa [15] provides a zoomable interface to multiple directories of images at once, grouping the images from the folders into clusters for maximal use of the screen. Harada et al. [16] proposed a browser focusing on automatically generated event assignments of photos; this method could well be implemented as a plug-in for our prototype. Girgensohn et al. [17] provide an interface for grouping and browsing photos based on several similarity measures: visual; geological; date; and tag. The Camelis photo browser uses co-occurrences of tags in images to deduce relationships and uses those relationships to facilitate browsing [18]. Several researchers have considered photo spreadsheets; in one of the most recent works, Kandel et al. focus on a biological application [19]. A good summary of the issues and early developments is found in [20].

As mentioned in the introduction, some efforts have been made towards adapting the MDA model to photo browsing [1–3]. These systems, however, are static representations of the respective collections, and hence ill-suited for general and personal photo browsing. The approach most similar to the MDA approach is that of *faceted search* which has been applied to many domains, including photos [4]. Faceted search uses a single set of tags, but proposes to build multiple hierarchies (or even DAGs) over that tag-set, one for each aspect that could be browsed. The single tag-set is a limitation; although the hierarchies do help users somewhat to disambiguate the different uses of an ambiguous tag, it is more logical to place distinct tags in different concepts. Furthermore, faceted browsers typically employ a linear presentation, resulting a dimensionality reduction of sorts, where it is unclear why photos appear on screen. Girgensohn et al. [21] propose a system that groups photos along multiple hierarchies, corresponding to different concepts, allowing each photo to appear logically in many places. By selecting different parts of different hierarchies, filters are applied to select subsets of the photos. The resulting photo set is still presented as single-dimensional, and each hierarchy can only contribute one filter to the set.

Scenique [5] is a faceted photo browser that breaks from tradition; it is conceptually the browser most similar to our proposal. Scenique allows image browsing in 3D browsing rooms, where each dimension corresponds to a facet. In addition to tag-based facets, Scenique also offers facets based on content-based descriptors. The M^3 model does not prevent this in any way and we plan to add support for content-based browsing to the server prototype. Scenique can not, however, show different parts of the same hierarchy of different browsing dimensions and in fact it is not clear how it handles the case when constraints are given for 1–2 or 4+ dimensions. We believe that the underlying abstractions of the M^3 model, in particular the cell and the hypercube, are the key difference.

5 Conclusion

Collections of digital media are growing ever larger, leaving users overwhelmed with data but lacking insights. Looking at the literature, the interest in photo browsing peaked shortly after the turn of the century and today fewer papers are published on photo browsing. Clearly, however, it remains an unsolved problem and one that merits further work. We have therefore proposed the M^3 data model for media browsing, based on the highly successful multi-dimensional analysis model from the business intelligence community, which is a natural progression of earlier work on faceted search. We have presented a detailed browsing scenario showing the expressiveness of the data model, as well as the results from two preliminary user studies which indicate that users find the data model and prototype both useful and engaging.

There are many interesting avenues for future work. Further development of the interface is necessary, of course, as are further user studies. We plan to add support similarity metrics, e.g. based on visual similarity measures, and for dynamic browsing dimensions, e.g., based on key-word search or content-based similarity, thus integrating browsing and searching into a single framework. Extending the model and prototypes towards professional and social—big—collections will also be interesting and challenging, both in terms of data model expressiveness and not least in terms of scalability. Furthermore, adapting to haptic and tactile interface is an interesting research direction.

Acknowledgments: This work was partially supported by Icelandic Research Fund grant 70005021, Student Innovation Fund grants 110395-0091 and 131956-0091, and CNRS PICS grant 6382, “MM-Analytics”.

References

1. Zaïane, O.R., Han, J., Li, Z.N., Hou, J.: Mining multimedia data. In: Proc. CASCON. (1998)
2. Arigon, A.M., Miquel, M., Tchounikine, A.: Multimedia data warehouses: A multiversion model and a medical application. Multimedia Tools and Applications **35**(1) (2007)
3. Jin, X., Han, J., Cao, L., Luo, J., Ding, B., Lin, C.X.: Visual cube and on-line analytical processing of images. In: Proc. CIKM. (2010)
4. Yee, K.P., Swearingen, K., Li, K., Hearst, M.: Faceted metadata for image search and browsing. In: Proc. CHI. (2003)
5. Bartolini, I., Ciaccia, P.: Integrating semantic and visual facets for browsing digital photo collections. In: Proc. SEBD. (2009)
6. Diao, M., Mukherjea, S., Rajput, N., Srivastava, K.: Faceted search and browsing of audio content on spoken web. In: Proc. CIKM. (2010)
7. Tómasson, G.: ObjectCube – a generic multi-dimensional model for media browsing. Master's thesis, Reykjavik University (2011)
8. Rúnarsson, K.: A face recognition plug-in for the PhotoCube browser. Master's thesis, Reykjavik University (2011)
9. Sigurþórsson, H.: PhotoCube: Multi-dimensional image browsing. Master's thesis, Reykjavik University (2011)
10. Tómasson, G., Sigurþórsson, H., Jónsson, B.Þ., Amsaleg, L.: Photocube: Effective and efficient multi-dimensional browsing of personal photo collections (demonstration). In: Proc. ICMR. (2011)
11. Tómasson, G., Sigurþórsson, H., Rúnarsson, K., Ólafsson, G.K., Jónsson, B.Þ., Amsaleg, L.: Photocube: Effective and efficient multi-dimensional browsing of personal photo collections (demonstration). In: Proc. ICMR. (2011)
12. Jónsson, B.Þ., Eiríksdóttir, A., Waage, O., Tómasson, G., Sigurþórsson, H., Amsaleg, L.: Photocube: Effective and efficient multi-dimensional browsing of personal photo collections (demonstration). In: Proc. ICMR. (2011)
13. Mathes, A.: Folksonomies - cooperative classification and communication through shared metadata. online (2004)
14. Kang, H., Shneiderman, B.: Visualization methods for personal photo collections: Browsing and searching in the photofinder. In: Proc. IEEE ICME (III). (2000)
15. Bederson, B.B.: PhotoMesa: A zoomable image browser using quantum treemaps and bubblemaps. In: Proc. UIST. (2001)
16. Harada, S., Naaman, M., Song, Y.J., Wang, Q., Paepcke, A.: Lost in memories: Interacting with large photo collections on PDAs. In: Proc. JCDL. (2004)
17. Girgensohn, A., Shipman, F., Turner, T., Wilcox, L.: Flexible access to photo libraries via time, place, tags, and visual features. In: Proc. JCDL. (2010)
18. Ferré, S.: Camelis: Organizing and browsing a personal photo collection with a logical information system. In: Proc. CLA. (2007)
19. Kandel, S., Paepcke, A., Theobald, M., Garcia-Molina, H., Abelson, E.: Photospread: A spreadsheet for managing photos. In: Proc. CHI. (2008)
20. Shneiderman, B., Bederson, B.B., Drucker, S.M.: Find that photo! interface strategies to annotate, browse, and share. Communications of the ACM **49**(4) (2006)
21. Girgensohn, A., Adcock, J., Cooper, M.L., Foote, J., Wilcox, L.: Simplifying the management of large photo collections. In: Proc. INTERACT. (2003)