Phorigami: Visualization of Digital Photo Collections by Origami Arts

Shuo-Hsiu Hsu[1], Pierre Cubaud[2], Sylvie Jumpertz[1],

[1] Orange Labs, 4, rue de Clos Courtel BP91226,
35512 Cesson Sévigné Cedex, France
[2] CNAM-CEDRIC, 55 rue du Turbigo, 75003 Paris, France
{shuohsiu.hsu, sylvie.jumpertz}@orange-ftgroup.com, cubaud@cnam.fr

Abstract. Issues about navigation of large photo collections have been studied for years. Associated interfaces tend to provide a single usage to deal with all kinds of photos and an integrated interface is rarely addressed. As the everyman photo collection consists of photos of various contexts, we advocate browsing different types of photos by different interaction techniques and presentation models. We present Phorigami, a photo browser based on proposed meta-categorization to enable searching by browsing. We describe the mapping between the six categories and the origami arts in terms of examining paper models with different levels of folding and the potential benefits for end-user experience is discussed in the end.

Keywords: Visualization, origami arts, personal photo collections, photo browsing.

1 Introduction

For people in this age of digital explosion, building digital personal photo collections becomes indispensable due to the availability of image capturing devices, storage equipments and various internet services. The drastic growth of personal digital photos published on the online sharing services, i.e., flickr, Google Picassa or saved in personal storage devices becomes a "black hole" from the viewpoint of management. Such growth of digital contents has consequently stressed the conventional, WIMP-based, management interfaces. Organizing the digital collections through categories or themes is an evident way of reducing or filtering the information flux for users. A posteriori categorization via content analysis techniques is a very promising approach, since the production of metadata is time consuming and inefficient for large public applications. As we have described from the work of Rodden [17], one of the interface requirements is to apply task-free techniques to facilitate the photo browsing. Therefore, the task-free management techniques such as the automatic photo clustering are heavily applied to the system tools. Such techniques exploit the add-on camera metadata to support different photo clustering strategies.

2 State of the art

Information visualization systems help users to understand data by presenting the results of data analysis. The core manner of organizing digital photo collections relies on the mechanism of categorization. A categorized content can facilitate browsing, especially when users have a clear target. Automatic algorithms for photo management are known as time-based [3] [4], location-based [5], content-based [6],[7] and multiple-clustering algorithms. These algorithms focus either on visual content of photos or on its metadata to perform clustering tasks according to the user's demand. As Rodden [8] indicates that a task-free management technique such as automatic photo clustering is one of the requirements for photo management system, researchers make the effort to explore more complex algorithms to satisfy varied demand from users.

Rodden et. al [6] apply the low level content-based analysis to conduct the image clustering by similarity; however, photo clustering by similarity may not enhance the efficiency in browsing when users do not have a clear idea about the target for their search browsing. PhotoTOC [22] is integrated with time-based and content-based clustering to automatically organize digital photos according to their events. The result of their experiment confirms positive feedback about using an automatic organization technique for the management of digital photo collection. Besides, FotoFile [21] and EasyAlbum [23] are applied with an annotation and content-based analysis technique to manage people by face recognition. Schaffalitzky et al. [19] investigate with content-based image analysis to establish a relative viewpoint from unordered digital photos. In addition, Jaimes et al. [18] raise the issue of detecting non-identical duplicate photos in consumer photo collections. They propose a content-based analysis algorithm and conduct an experiment by manual classification to observe the performance of proposed algorithm; however no concrete interface is addressed in their study. All in all, the conventional management tools with automatic clustering techniques for the end user, to this day, are only used to deal with the demand in a specific situation. The relevance of these specific situations with respect to "everyman" photo collection is rarely addressed in the literature.

Photowares [16] based on content analysis techniques mainly focus on enhancing its performance of specifying queries: however their visualization and interaction techniques still rely on the traditional WIMP interface. The deficiency of traditional WIMP interface for those InfoVis tools is their limited and fixed visualization techniques that may not sufficiently support various tasks of browsing large collections. Although certain interface studies such as PhotoMesa [1], Photo tourism [20], Face Bubble [24] and the work of Porta [14] have tried to make up for this shortcoming by integrating unconventional visualization and interaction techniques into experimental photo browsers, however the visualization and interaction techniques beneficial for specific usages are still in search.

2.1 Overview of visualization techniques

The objective of visualization techniques is to feature the prominent part of the analyzed data. For years, several techniques have been proposed for different problematics of visualizing large data. The most common visualization technique for large content data is the thumbnail display that resized data units providing users with a full view of the content. However, for visual contents such as images, the drawback of a thumbnail display is the loss of informativity and visibility with the resized thumbnail presentation. The zooming techniques are therefore developed to improve the informativity and visibility of the content. Certain context+focus techniques such as fisheye, sphere visualization [9] and hypertree browser [2] allow users to keep both foreground and sideward information when zooming into a focused part. However, both fisheye and sphere effects may cause the loss of partial visual information due to image distortion. Another advanced version of the zooming technique is the zoomable interface integrated with a treemap such as PhotoMesa [1]. With the same context-focused technique, a zoomable browsing interface allows users to be less disoriented when browsing categorized data. Other context+focus techniques applied in visualizing large data can be found as perspective wall [10] and melange [11] which both apply the folding technique to make the focused part stand out. Moreover, development of visualization techniques also goes to another way of different presentation modes. The Rapid Serial Visual Presentation, RSVP [12] aims to provide an automatic presentation so that users can browse more content within a short time. The study of [13] compares different RSVP modes for photo presentation, and Porta [14] proposed varied artistic visualization techniques based on RSVP for presenting large image collections. In spite of the advantage of automatic presentation, the crucial drawback is that the fixed presentation speed (no matter fast or slow) fails to satisfy the realistic human behavior of browsing.

3 Categorization of photo collections

The drawback of current studies in automatic categorization techniques is the lack of an integrated interface to deal with varied types of photos for our everyman albums. Thus we present Phorigami, a photo browser based on the meta-categorization and origami visualization. This categorization approach encompasses the scope of current or expected recognition technologies. Our goal is not only to propose a categorization approach, but also to outline different interaction and presentation models toward different categories. Concerning the meta-categorization, the details about the experiments and the discussion of findings are presented in the associated article [29]. We then outline our meta-interface by applying different interaction techniques and origami visualization to feature each photo group.

3.1 Meta-categorization for everyman photo collections

We propose a meta-categorization method based on the analysis of the relative status between the camera and the focused targets. The objective of categorization is to

perform an automatic organization of photo contents and to highlight different contexts in the photo collection. This proposed approach can potentially be implemented by the state of the art of content-based image analysis algorithm and camera metadata. We analyze the photo shooting in terms of two parameters: the movement of the camera and the focused targets in two states (static and mobile). A similar idea of analyzing photos by camera and photographer has been used in Jaimes et al. [15] in terms of camera, scene and image parameters to detect duplicate digital photos. In our definition, we add a third dimension called "groups" for targets, in order to include the user experience in taking group photos of reunion. The Table 1 presents the distribution of parameters and associated dimensions and six classifications are generated as A, B, C, D, E, F.

Table 1. Movement of camera and focused targets.

Target of photographer \ Camera	Fixed	Mobile
Fixed	**A**	**B**
Mobile	**C**	**D**
Groups(subject replacement)	**E**	**F**

Table 2. Classifications with associated contexts and interaction techniques.

Type	Intention of Photographer	Photo Context	Interaction Technique
A	Simple static view	Panorama	Panoramic Presentation
B	Multi-view		
C	Motion capturing	Action	Animated Photo Presentation
D	Motion capturing		
E	Groups	Social Relation	Simple Folding Presentation
F	Groups in motion and subject replacement		

We explain each classification with associated photo scenarios in Table 3. Type A refers to a scenario of a simple static view point where each single photo is focused on a static object from the same environment; for example, a user takes photos of several artworks on an exhibition and the scenario of "multi-view" in type B describes how the photographer changes the position to take a photo of each face of a static object, especially for buildings. Associated examples are illustrated in Fig. 1. Types C and D belong to the motion-capturing scenario where the photographer takes photos of a moving object. The difference between C and D depends on the status of the camera. An example of type C is presented in Fig. 2 where the photographer stays motionless to take photos of walking pedestrians with a bird's-eye view. Type D shows where the photographer moves the camera to trace an object in motion. Likewise, types E and F in Fig. 3 are the group photos of reunion. It is to be noted that type E refers to photos of a static group while photos concerning the subject replacement [15] or changes in movement are classified as type F. In addition, the double-shot or multi-shot of a digital camera on the same target are what have caused

the problems in managing digital photo collections [15]. In our categorization, such double-shot photos in the same context will be possibly categorized in type B, C, D, F while the intention of the photographer is merely to take a "better" photo.

Fig. 1. Type A: Simple static view(left), Type B: Multi-view (right)

Fig. 2. Type C: Motion capturing with fixed camera (left), Type D: Motion capturing with mobile camera (right)

Fig. 3. Type E: Groups (left), Type F: Groups in motion(right)

3.2 Origami visualization based on meta-categorization

We explore appropriate metaphorical mappings for our interaction techniques in terms of two directions: human experience in interacting with paper-based documents and the art of origami. When browsing a large photo collection, it is important to provide a global view of the entire content and at the same time allow users to be information-conscious toward each category. The conventional thumbnail display succeeds at providing a global view for browsing the large content; however it is space-consuming and may cause visual overloads. As we aim to feature the photo context by applying our categorization approach, each photo context is supposed to be presented by a specific visualization technique. Such visualization technique requires the visual capacity to embody the photo context via the appropriate metaphorical mapping and also requires the interaction capacity to provide users an intuitive manipulative technique.

The principle of the folding technique in origami art is to reconstruct an object via a minimal presentation. The folding technique is implemented from an everyday use to an artistic presentation. The brochure is one of the practical usages of the folding technique in the daily life (see Fig.4). The law of folding in origami art has been applied to solve problems in engineering, industrial design and scientific work [25]. The commercial products implemented with folding techniques can also be found in

Rovi Liquid media guide interface launched in 2009 [26] and Scentsory-Nokia mobile phone concept [27]. The folding visualization in Rovi Liquid media guide interface tends to give users a continuous presentation for associated TV programs and multimedia contents. For the concept design of Nokia mobile phone, the objective is to enhance the mobility of the device by the folding mechanism and also to provide an innovative user interface. In the literature, the folding technique is firstly applied as a context-focus technique in Perspective wall [10] that a simple folding technique helps to feature the focused information. Similarly, the folding technique is proposed in "mélange" [11] for multi-focus interaction in a complex InfoVis system.

The goal of this folding model is to examine the capacity of the folding presentation to convey photos. We aim at the fast prototyping paper models to acquire appropriate mapping between the folding function and the photo visualization of our six categories in three contexts. Prototypes are done with folding in three levels of complexity: low, medium and high. In Fig.5, we present certain snapshots about the development of folding ideas. We build a 4x4 block square with two themes of photos (theme of a cat and theme of trips) printed on the different side. This folding mechanism expands photos from the centre of the object and its presentation may probably correspond to the metaphor of "groups getting together". We make another folding model whose feature is minimizing the object by a vertically-extended folding mechanisms. The original model and the photo-printed model are shown in Fig.6. We find that printed photos on the folded paper are distorted and less visible for two reasons. First, photos are less informative because they are distorted by complex folds. Second, the vertically extended folding fails to present the photo contexts in a consistent way. We consider simplifying this folding mechanism and apply it to present a series of snapshot photos in our meta-interface. Another trial of the folding paper model is shown in Fig.7 where the surface of the paper is entirely folded so that the folded presentation can be more freely manipulated. However, this folding idea fails to contribute to one of our three photo contexts due to its high freedom in folding.

For the photo context of “action”, the best way to browse such kind of photos is to allow them to play as a filmstrip. The filmstrip view is broadly implemented in conventional software interfaces especially for audiovisual software. The feature of the filmstrip view is to provide a browsing window for the ongoing displayed content and a linear window to present the entire image frames; however its drawback is the requirement of sufficient space for two window displays. We trace back to the early invention used to project motion pictures. One of the early devices for displaying motion pictures is Zoopraxiscope created by Eadwear Muybridge in 1879 [28]. The image frames are printed on a glass disk, which is equal to the function of videotape. For our meta-interface, the round layout may be space consuming so that an appropriate layout for displaying the motion pictures requires to be taken into account.

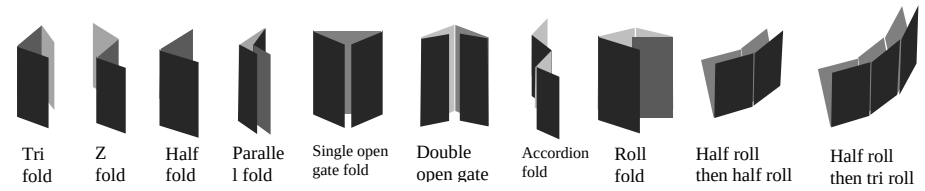

Fig.4. Different type of brochures.

Fig.5. Folding in low complexity. Unfolded (left), half-folded (middle) and folded (right).

Fig.6. Folding in medium complexity: original paper model (left two), paper model with printed photos (right two).

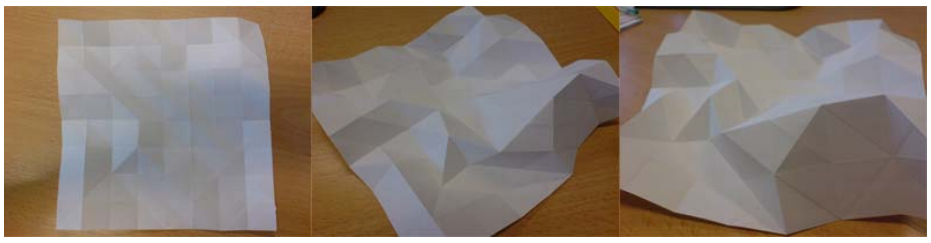

Fig.7. Folding in high complexity.

3.3 Phorigami meta-interface

Based on the idea development described above, we outline our interface, Phorigami, with different origami visualization: Accordion folding display, Rolodex display and a simple folding display, as illustrated in Fig.8.

For photos of Category A and B, we use an accordion-like folding technique to compress the space of panoramic photos into a folding mode to present an entire panoramic photo by a horizontal photo extension. The mechanism of accordion fold corresponds to the continuous photo snapshots toward a panoramic scene. For motion frames in Category C and D, the potential mechanism is supposed to manipulate motion frames like a GIF animation. Same as the automatic presentation, the related technique such as RSVP specializes in presenting images with a specific speed; however, such preset browsing speed fails to be adapted in different contexts and may counteract the user behavior in browsing images. Therefore, we propose a "manual" Rolodex technique that users can customize the browsing speed to render the original scenario via motion photos. In addition, we simplify and refine the previously-developed idea of 4x4 block square. Due to the consideration of the layout, a center-expanded presentation may be difficult to unfold when its position is near the boundary layout. Thus a poker pile presentation is used to visualize photos of Category E and F.

It is to be noted, for a folded status, the size of each presentation model except the accordion presentation is maintained within a single square size as a thumbnail image. For poker pile and Rolodex presentation, users can learn to understand their visual representation both from the thumbnail image and the shape of presentation; however the accordion presentation fails to provide clear thumbnail image because its image content is folded. Therefore, the basic size of the folded accordion presentation is defined as two squares in order to reveal more visual messages for users to understand the photo content from the folded presentation.

We develop our interface prototype by using the Processing environment to simulate the visualization and associated interaction techniques. The three visualization models are supposed to be implemented with tactile technology to enhance the sense of manipulation toward virtual objects. As illustrated in Fig.8, users are allowed to interact with three presentation models by finger dragging or nudging. In Fig.9, we present the snapshot of Phorigami visualizing 561 photos. As illustrated, the screen space used by Phorigami is more than three times less than that used by a conventional thumbnail interface. In general, photo albums presented by Phorigami potentially present rich visualization results, which are both information-conscious and space-saving.

The potential benefits of Phorigami to the end-user experience can be supposed in two aspects: first, it may trigger the visual clues beneficial to conduct different levels of browsing, i.e., general browsing, search browsing and serendipitous browsing[1]. Second, the meta-interface may enhance the pleasure of browsing due to providing

[1] Rice, R.E., McCreadie, M.M. and Change, S.L. (2001), Accessing and Browsing Information and Communication, MIT Press, Cambridge, MA. p.182.

different interaction techniques.

This prototype has been presented informally to obtain the preliminary user feedback. In general, users confirm the coherence between the categorized photos and presented visualization techniques. The accordion and Rolodex presentation are much more fascinated than the poker pile one. However, the view of folded content seems to be somehow disordered that the three presentation models fail to be highlighted and visually interfered by other uncategorized photos. The prototype may require the application of a further visual effect to rectify this deficiency. The potential solution may be to assign the categorized and uncategorized photos in different layers in order to distinguish the difference. As the formal process, the usability of the different interaction techniques and related metaphor mapping require an integral test plan to execute associated tasks to examine our interface prototype.

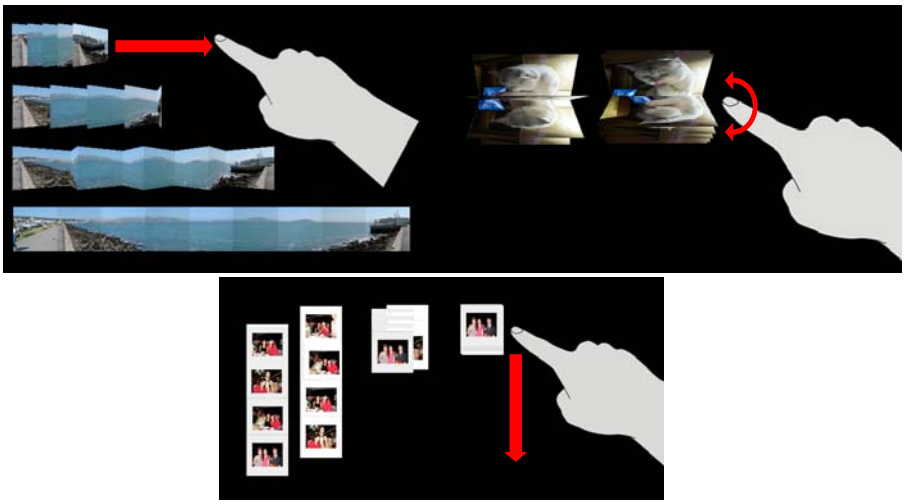

Fig.8. Accordion presentation (top left), Poker pile presentation (bottom), Rolodex presentation (top right) with dragging and nudging.

4 Conclusion

For years, the photo management system tools have evolved into more powerful algorithms that enhance the efficiency in dealing with the visual content of digital images. Different from the development based on a technology-oriented viewpoint, we come to deal with the management of digital photos from a more user-centered standpoint, envisioning a potential interface by proposing a framework for the future system.

Fig.9. Visualization of 561 photos by Phorigami. The folded content view(Top) and the original content view (Bottom).

For the digital photo management, our original idea is evoked from the user experience on daily practice of photo shooting. This idea is enriched and expanded into a categorization approach based on the state of art technology. Our goal is far from a simple categorization solution but opens a vision with a user-centered interface design to facilitate the organization of digital photos. While conventional approaches are used to working on add-on techniques to facilitate photo categorization, we aim to stand out different photo contexts for photo categorization from a more user-centered viewpoint. Although nowaday system tools support photo categorization by providing diverse techniques, i.e. text tagging, time clustering and content-based analysis, we advocate conducting photo categorization via their contexts. The photo categorization data generated by clustering algorithms is entirely a posteriori while the photo context can be discerned at a glance without any added-on processing. The photo context as the complex information can be narrated in various ways by textual description but can easily be understood by visual presentations. Moreover, the photo context representing a variety of emotional clues can visually recall user's memory to render associated scenarios and stories. All in all, we confirm the importance of state of the art development in clustering algorithms and consider our user-centered interface proposal to be a potential framework to develop a future system for digital photo management. The future work will aim at the user studies on comparing Phorigami with conventional thumbnail browser, especially on the issue of serendipitous browsing.

Acknowledgments. We would like to thank Rodrigo A. B. Almeida for his opinions and the help on the development of the prototype in Processing programming. We are grateful to the anonymous referee for her/his time and efforts in providing the constructive suggestions. Thanks also go to Chia-yi Shy for the sharing of her digital photo albums as the sample collection for the demonstration of our interface design.

References

1. Bederson, B.B. PhotoMesa: A Zoomable Image Browser using quantum treemaps and bubblemaps. ACM UIST'2001, pp. 71-80 (2001)
2.Lamping, J., Rao, R., Pirolli, P.: A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. Conference on Human Factors in Computing Systems, pp. 401-408 (1995)
3. Graham, A., Garcia-Molina, H., Paepcke, A., Winograd, T.: Time as essence for photo browsing through personal digital libraries. Joint Conference on Digital Libraries, pp. 326-335 (2002)
4. Harada, S., Naaman, M., Song, Y. J., Wang, Q. Y., Paepcke, A.: Lost in memories: interacting with photo collections on PDAs. Joint Conference on Digital Libraries, pp. 325-333 (2004)
5. Toyama, K., Logan, R., Roseway, A.: Geographic location tags on digital images. ACM Multimedia, pp.156-166 (2003)
6. Rodden, K., Basalaj, W., Sinclair, D., Wood, K., R.: Does Organization by Similarity Assist Image browsing? Conference on Human Factors in Computing Systems, pp.190-197 (2001)

7. Xiao, J., Zhang, T.: Face Bubble: Photo Browsing by Faces. International Working Conference on Advanced Visual Interfaces, pp. 343-346 (2008)
8.Rodden, K. How Do People Organise Their Photographs? BCS-IRSG Annual Colloquium on IR Research (1999)
9.Christmann, O., Carbonell, N.: Browsing through 3D Representations of Unstructured Picture Collections: an Empirical Study. International Working Conference on Advanced Visual Interfaces, pp.445-448 (2006)
10.Mackinlay, J., D., Robertson, G., G., Card, S., K.: The perspective wall: detail and context smoothly integrated. CHI, pp. 173-176 (1991)
11.Elmqvist, N., Henry, N., Riche, Y., Fekete, J., D.: Melange: Spcae folding for Multi-focus Interaction. Conference on Human Factors in Computing Systems, pp. 1333-1342 (2008)
12.Spence, R.: Rapid, Serial and Visual: a presentation technique with potential. Information Visualization 1(1), pp.13-19 (2002)
13.Cooper, K., Bruijn, O., Spence, R., Witkowski, M.: A Comparison of Static and Moving Presentation Modes for Image Collections. International Working Conference on Advanced Visual Interfaces, pp. 381-388 (2006)
14.Porta, M.: Browsing Large Collections of Images through Unconventional Visualization Techniques. International Working Conference on Advanced Visual Interfaces, pp. 440-444 (2006)
15.Jaimes, A., Chang, S. F., Loui, A. Detection of non-identical duplicate consumer photographs. Pacific Rim conference on Multimedia, pp.16-20 (2003)
16.Kirk,D., Sellen,A., Rother, C., Wood, K.R. Understanding photowork. ACM CHI'2006, pp. 761-770(2006)
17.Rodden, K. How Do People Organise Their Photographs?. BCS-IRSG Annual Colloquium on IR Research (1999)
18.Jaimes, A., Chang, S. F., Loui, A. Detection of non-identical duplicate consumer photographs. Pacific Rim conference on Multimedia, pp.16-20 (2003)
19.Schaffalitzky, F., Zisserman, A. Multi-view Matching for Unordered Image Sets, or "How Do I Organize My Holiday Snaps?. ECCV'2002, pp. 414-431(2002)
20.Snavely, N., Seitz, S. M., Szeliski, R.. Photo tourism: exploring photo collections in 3D. ACM Trans. Graph. 25(3), pp.835-846 (2006)
21.Kuchinsky, A., Pering, C., Creech, M. L., Freeze, D. F., Serra, B., Gwizdka, J. *FotoFile*: A Consumer Multimedia Organization and Retrieval System. ACM CHI'1999, pp. 496-503(1999)
22.Platt, J., Czerwinski, M., Field, B. PhotoTOC: Automatic Clustering for Browsing Personal Photographs, Fourth IEEE Pacific Rim Conference on Multimedia. Vol.1, pp.6-10 (2003)
23.Cui, J., Wen, F., Xiao, R., Tian, Y., Tang, X. EasyAlbum: an interactive photo annotation system based on face clustering and re-ranking. ACM CHI'2007, pp.367-376(2007)
24.Xiao, J., Zhang, T. Face Bubble: photo browsing by faces. ACM AVI'2008, pp.343-346
25.Robert Lang. Idea + Square = Orgami
http://www.ted.com/index.php/talks/robert_lang_folds_way_new_origami.html
26.Rovi Liquid media guide interface http://news.cnet.com/8301-17938_105-10288475-1.html
27.Nokia scentsory, http://www.intomobile.com/2007/05/21/nokia-scentsory-cellphone-concept.html
28.Zoopraxiscope http://en.wikipedia.org/wiki/Zoopraxiscope
29.Hsu, S.H., Cubaud, P., Jumpertz, S.: Phorigami: A Photo browser based on méta-categorization and origami visualization. International conference on human-computer interaction, San Diego, USA, pp.801-810(2009)