

# A dynamic flexible and interactive display method of digital photographs <sup>☆</sup>



Chi Thanh Vi <sup>a,\*</sup>, Kazuki Takashima <sup>b</sup>, Hitomi Yokoyama <sup>b,1</sup>, Gengdai Liu <sup>c</sup>, Yuichi Itoh <sup>d</sup>, Sriram Subramanian <sup>a</sup>, Yoshifumi Kitamura <sup>b</sup>

<sup>a</sup> Department of Computer Science, University of Bristol, Woodland Road, Bristol BS8 1UB, United Kingdom

<sup>b</sup> Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan

<sup>c</sup> OLM Digital Inc., Mikami Bldg, 2F, 1-18-10 Wakabayashi, Setagaya-ku, Tokyo 154-0023, Japan

<sup>d</sup> Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

## ARTICLE INFO

### Article history:

Received 10 April 2014

Revised 31 July 2014

Accepted 1 August 2014

Available online 14 August 2014

### Keywords:

Entertainment computing

Animation

Emergent computation

Photograph

Task engagement

Electroencephalography (EEG)

## ABSTRACT

We propose D-FLIP (Dynamic & Flexible Interactive PhotoShow), a novel algorithm that dynamically displays digital photos using different organizing principles. Various requirements for photo arrangements can be flexibly replaced or added through the interaction and the results are continuously and dynamically displayed. D-FLIP uses an approach based on combinatorial optimization and emergent computation, where geometric parameters such as location, size, and photo angle are considered to be functions of time; dynamically determined by local relationships among adjacent photos at every time instance. As a consequence, the global layout of all photos is automatically varied. We first present examples of photograph behaviors that demonstrate the algorithm and then investigate users' task engagement using EEG in the context of story preparation and telling. The result shows that D-FLIP requires less task engagement and mental efforts in order to support storytelling.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Pervasiveness of digital cameras has led to large collections of digital photos that user's often browse on computer displays, by rearranging them to gather similar ones based on specific features/meta-data. Although several techniques to do this efficiently exist, most of these are somewhat systematic or goal-driven in terms of applying principles for displaying photos. These methods are useful in systematically organizing and finding photos but previous studies suggest that users often browse their photo collections without a specific search goal (e.g. [1]). Moreover, users usually browse photos with actions such as enlarging displayed thumbnails in a certain order, displaying photos randomly on a digital photo frame or starting a slideshow for personal

gratification and pleasure. To support and enrich these photo-browsing behaviors, the presentation of photos should flexibly and dynamically adapt with visual effects based on user's input.

Additionally, one of the most enjoyable parts of personal photos is to share memories and reminisce with friends or relatives. Previous attempts to provide such experiences have revolved around presenting a static collection with interaction capabilities on tables [2] and handheld devices to facilitate story telling [3]. However, we want to improve not only the display layout but also the dynamic behaviors of the photos during interactions.

Therefore, we propose a novel method to flexibly display a set of photos by showing each of them in a dynamic and continuous motion like a living object. It allows users to replace or add displaying principles interactively and flexibly. In addition, users can manipulate (such as enlarging and translating) a particular photo through flexibly grouping and arranging them using meta-data and/or their feature extracted values. In order to achieve such flexibility, we introduce an approach based on emergent computation. Geometric parameters (i.e. location, size, and photo angle) are considered to be functions of time. Photos are dynamically moved toward the directions determined by local relationships with adjacent photos at each time instance. As a result, the global layout of all photos varies automatically; converging gradually with time.

<sup>☆</sup> This paper has been recommended for acceptance by Haruhiro Katayose.

\* Corresponding author.

E-mail addresses: [Chi.Vi@bristol.ac.uk](mailto:Chi.Vi@bristol.ac.uk) (C.T. Vi), [takashima@iec.tohoku.ac.jp](mailto:takashima@iec.tohoku.ac.jp) (K. Takashima), [hitomi-y@cc.tuat.ac.jp](mailto:hitomi-y@cc.tuat.ac.jp) (H. Yokoyama), [liugengdai@gmail.com](mailto:liugengdai@gmail.com) (G. Liu), [itoh@ist.osaka-u.ac.jp](mailto:itoh@ist.osaka-u.ac.jp) (Y. Itoh), [Sriram.Subramanian@bristol.ac.uk](mailto:Sriram.Subramanian@bristol.ac.uk) (S. Subramanian), [kitamura@iec.tohoku.ac.jp](mailto:kitamura@iec.tohoku.ac.jp) (Y. Kitamura).

<sup>1</sup> Tokyo University of Agriculture and Technology, Graduate School of Engineering, Advanced Information Sciences Division, 184-8588 Koganei-shi, Naka, Tokyo 2-24-16, Japan.

This will enhance one of the most enjoyable parts of personal photos, which is to share memories and reminisce with friends or relatives.

We illustrate example behaviors of photos and then do a user study to evaluate D-FLIP against Windows Explorer, a photo managing program familiar to Windows users. The evaluation involved two participants, a narrator and a listener to prepare and share a story. We measured both participants EEG signals to quantitatively measure users' mental effort/task engagement. Moreover, NASA-TLX forms were collected from the narrators and listeners after each task.

The contributions of this paper are: (1) a proposed method to dynamically and flexibly display photos; and (2) an evaluation method using EEG that can be used to evaluate interactive applications.

## 2. Related work

### 2.1. Browsing digital photos

Many efforts have been proposed to arrange photos effectively. For example, a browser that arranges multiple photos in folders by grouping them with different magnification levels [4], by categories with different hierarchy depths [5], or by displaying more photos in a meaningful way (e.g. adequate groupings [6]). Other examples are arranging photos calendar by using their shooting dates [7], displaying them on a digital geographical map at their shoot locations using meta-data [8], grouping photos with shoot locations and persons [9], and browsing large image datasets using Voronoi diagrams [10]. A technique for browsing large image collections was presented by [11] using the rectangle-packing algorithm, and by [12] using hierarchical tree structured organization of images with level of details (LOD). Additionally, [13] proposed an extended slide show in which multiple photos are sequentially displayed in a fixed tiling manner with appropriate music. However, most of these methods allow users to handle photos somewhat in a systematic manner, in terms of selection of requirements or principles of displaying photos. They lack flexibility in displaying with variety of requirements based on user's input.

### 2.2. Photo collages and combinatorial optimization

Digital photo collages, which summarize meaningful events or memorabilia, are widely used to display photos. This is efficient because users can view multiple photos at once. However, it requires two types of treatment: (1) a geometric treatment concerning about arranging multiple photos in a pre-determined area but avoids overlapping and empty regions as much as possible, and (2) a semantic treatment concerning about content of the photos. Several authoring tools have been proposed to create photo collages easily. For example, AutoCollage [14] is an algorithm to create a collage automatically with little duplication of photos with smoothly blended natural boundaries using a sequence of energy minimization steps. Similarly, Picture Collage [15] is an algorithm to create a collage that shows as many visible salient regions as possible using the Markov chain Monte Carlo method for optimization. Digital Tapestry [16] creates collages using Markov Random Field.

Generally, the quality of the collages (e.g. the arrangement beautifulness and avoidance of overlapping) is represented by an energy function. This approach aims to maximize (or minimize) the function by solving the combinatorial optimization. Meta heuristic approaches are usually used to find an optimum solution. However, it is difficult to try different photo arrangements by

adding or replacing principles flexibly. This is because the quality of a solution needs to be evaluated in every time the condition is varied.

### 2.3. Emerging computing and Force-based method

Emergent computing is often used to achieve hi-level dynamic behaviors of individuals with simple local rules. Typical examples are: flock of birds [17], and a social force model to measure internal motivations of individual pedestrians [18]. Since then, the crowd simulation has been actively studied with many researches such as boids model to visualize time-varying data [19], or simulation of the formation process of spots and stripes on the animal's skin using the concept of emergent algorithm [20].

Multiple reaction-diffusion systems can be cascaded using iteration loops to create complex patterns by considering a simple repulsive force (which drops off linearly with distances) to push points. Shimada and Gossard [21] proposed an algorithm to generate triangular meshes using a Voronoi diagram by defining repulsive and attractive forces to perform dynamic simulation for a force-balancing configuration. Besides, force-based relaxation algorithms have been widely applied to the problems in constraint satisfaction with local propagation, such as graph drawing [22], information visualization [23], layout of objects [24], and crowd simulation [18], and so on.

Our method is based on methodologies of these categories, but more flexibility is achieved in order to replace or add principles of displaying photos.

### 2.4. The effect of animation on users' interest

Compared to static method, dynamic photo displaying seems to be more interesting and aesthetically appealing [25]. Previous studies have shown that animation can boost users' performance in learning and teaching such as understanding Newton's law of motion [26]. In other words, animations can help users perform the task (e.g. learning) easier and with a better performance. In terms of users' interest, both static and animated graphics can increase interest. Animation are likely to increase emotional interest (created by events that are arousing) while static graphics are likely to trigger more cognitive interest (related to the connections between incoming information and background understanding [27]). As the result, D-FLIP is likely to trigger emotional interest from users because of its dynamical and interactive movements. This will help to achieve the goal of D-FLIP that is letting users viewing photos interactively with ease and interest.

### 2.5. Evaluating using neural signals

Traditionally interactive programs are evaluated by investigating performance (i.e. [28]), or users' behaviors. However, with programs designed for using with ease and pleasure, an evaluation method measuring users' affective or inner states is more suitable. Usually this is done by questionnaires answered by participants during the experiment. However, they occur after the event when important issues may be forgotten. Additionally participants may not be aware of their states or might simply guess. Neural signals, measured from the brain can better reflect a users' current state and provide an evaluation metric.

There are different methods to detect neural signals such as fMRI, MEG, fNIRS and EEG. A brief summary of those techniques are discussed in [29]. In addition, EEG devices are portable (compared to fMRI, MEG) and have high temporal resolution (compared to fNIRS). EEG signals have been also shown to capture the affective state (such as relaxation [30], arousal [31], and task engagement [32]).

One purpose of our proposed algorithm, D-FLIP, is to help user browse photos with ease and interest thus measuring task engagement can help to evaluate. According to Matthews et al. [33], task engagement can be defined as the effortful concentration and striving towards task goals where task demands and personal characteristics may influence this pattern of processing. In another work, Matthews et al. noted that effort investments in the task would increase task engagement [34]. Studies have shown a positive correlation between EEG engagement and task demands including stimulus complexity processing and the requirement for attentional resources allocation [35,36]. EEG engagement correlated with task loads in not only simple vigilance and memory tasks but also in complex simulation tasks such as in radar operations simulation environment [35,37]. Consequently, if an application requires low level of task engagement in using, it also requires low level of task demands, task loads, and workloads. In other words, it is easier to use when compared with applications requiring higher task engagement.

Pope et al. [38] present a measurement of task engagement from EEG as  $\beta/(\alpha + \theta)$ . Further studies [39,40], with the same measure were used to show the performance quality benefits and reduced mental workload. These positive findings have been also replicated using extended periods of task performance [41] and a vigilance task [42]. Given this evidence of measuring users' task engagement/workload, we used this formula to evaluate our system in comparing D-FLIP with a similar and competitive program.

### 3. Algorithm overview of dynamic display of photos

Each photo has three parameters: its position, size, and rotational angle. These parameters are considered as functions of time and are controlled to arrange multiple photos simultaneously on a display. The photo movement is shown by gradually changing the values of these parameters at every time instant.

The algorithm is explained by Eq. (1):

$$d\vec{x}/dt = f(\vec{x}) + \eta \quad (1)$$

Here,  $\vec{x}$  is a set of the three parameters above and its variation  $d\vec{x}/dt$  is derived by  $f(\vec{x})$ , the principle to achieve the photo arrangement, and noise term  $\eta$ . Larger amplitude noise increases the fluctuation and is useful for escaping local optimum.

Furthermore, Eq. (1) can be re-written in another form with the weight coefficients:

$$\frac{d\vec{x}}{dt} = \sum_i \{w_i f_i(\vec{x}) + \eta_i\} \quad (2)$$

In here,  $f(\vec{x})$ , a variety of principles, is used to achieve the photos arrangement or layout. Assuming:

$P$	Data of a photo
$I$	Information of certain input or output devices
$\vec{P}$	All the photos in the environment
$Position(P)$	Photo position
$Size(P)$	Photo size
$Rotation(P)$	Photo rotational angle
$l, m, n$	The number of principles related to position, size, and rotation angle, respectively

Then Eq. (3) is obtained by modifying Eq. (2). It controls the parameters of photo  $P$  and is calculated from all photos. Here,  $f_{Pi}(\vec{x})$ ,  $f_{Si}(\vec{x})$  and  $f_{Ri}(\vec{x})$  are functions that represent the changes of position, size, and rotation, respectively:

$$\begin{cases} \frac{d}{dt} Position(\vec{P}) = \sum_i^n \{f_{Pi}(I, \vec{P}) + \eta_i\} \\ \frac{d}{dt} Scale(\vec{P}) = \sum_i^m \{f_{Si}(I, \vec{P}) + \eta_i\} \\ \frac{d}{dt} Rotation(\vec{P}) = \sum_i^l \{f_{Ri}(I, \vec{P}) + \eta_i\} \end{cases} \quad (3)$$

## 4. Principles of photograph arrangement

There are two types of principles that are important for photo arrangement: packing and mapping. Packing is a geometric problem concerning about arranging multiple photos with different sizes and rotational angles in a pre-determined area; it avoids overlaps and empty regions as much as possible. On the other hand, mapping is a semantic concerning about locating each photo based on its content and interaction with users. One example of packing is the function to enlarge each photo as much as possible, but avoid overlaps with adjacent photos by translation, rotation, or shrinking. Another example is the function to move photos toward the inside the displaying window to avoid exceeding its boundary. Examples of mapping are functions that attract photos with the same tag, or to enlarge an interesting photo by a viewer, and so on. Here, each function can be established independently based on an individual principle as well as to be implemented without paying attention to the global coordination. Certain feature values of each photo are assumed to be calculated and stored in the tag beforehand (e.g. to specify a person, taken location, etc.). By replacing or adding functions that correspond to the displaying principles, different photo arrangements can be achieved flexibly.

### 4.1. Geometric packing

Here we explain related principles related to geometric packing. First, the principle to avoid overlaps with adjacent photos is represented by Eq. (4). Here,  $N$  is the number of photos,  $Avoid(P, P_i)$  is  $P$ 's vector for escaping when  $P$  and  $P_i$  overlap.  $Adjacency(P)$  is the set of photos overlapping with  $P$ .

$$f_{translation}(I, \vec{P}) = \sum_i^N Avoid(P, P_i) \quad \text{if } P_i \in Adjacency(P) \quad (4)$$

Second, a photo is moved toward the inside of the window based on Eq. (5) if its position exceeds the displaying window's border. Here,  $L, B, R,$  and  $T$  are the *left, bottom, right,* and *top* coordinates of the window,  $L(P), B(P), R(P),$  and  $T(P)$  are the corresponding photo coordinates, and  $A_l, A_b, A_r,$  and  $A_t$  are their coefficients, respectively:

$$f_{mold}(I, \vec{P}) = \sum_i^N \left\{ \begin{array}{ll} A_l \{L - L(P_i)\} & \text{if } L(P_i) < L \\ A_b \{B - B(P_i)\} & \text{if } B(P_i) < B \\ A_r \{R - R(P_i)\} & \text{if } R(P_i) > R \\ A_t \{T - T(P_i)\} & \text{if } T(P_i) > T \end{array} \right\} \quad (5)$$

Fig. 1 illustrates how photos avoid overlapping. Without overlaps, each photo becomes larger until it reaches the predetermined maximum scale when Eq. (6) is applied (Fig. 1a). If two adjacent photos overlap, the larger photo becomes smaller until it reaches the predetermined minimum scale when Eq. (7) is applied (Fig. 1b); they move to opposite directions when Eq. (5) is applied (Fig. 1a), or rotate in opposite directions when Eq. (8) is applied (Fig. 1c). Here,  $A_{s1}$  and  $A_{s2}$  are coefficients, and  $Ang(P_i, P_j)$  is the rotational angle with which  $P_i$  and  $P_j$  avoid overlapping:

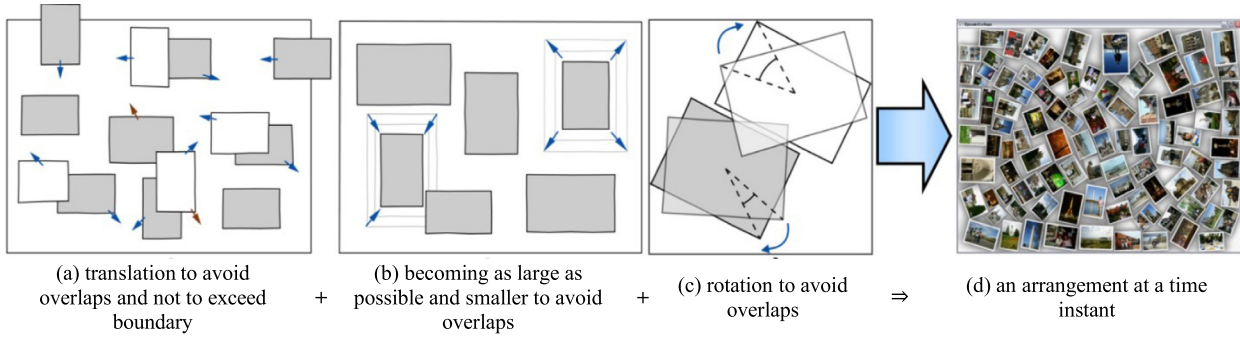


Fig. 1. Conceptual behaviors of photographs.

$$f_{enlarge}(I, \vec{P}) = \sum_i^N A_{s2} \{Scale_{max} - Scale(P_i)\} \text{ if } Adjacency(P_i) \\ = \varphi \wedge Scale_{max} > Scale(P_i) \quad (6)$$

$$f_{shrink}(I, \vec{P}) = \sum_i^N A_{s1} \{Scale_{min} - Scale(P_i)\} \text{ for all } P_j \\ \in Adjacency(P_i), Scale(P_j) < Scale(P_i) \wedge Scale_{min} \\ < Scale(P_i) \quad (7)$$

$$f_{rotation}(I, \vec{P}) = \sum_i^N Ang(P, P_i) \quad (8)$$

The upper-right photo in Fig. 1b will become as large as possible by referring to environmental parameters indicating the positions and sizes of adjacent photos. However, when two photos collide, the larger one becomes smaller, as shown in the lower right photo based on Eq. (7) if these two equations are simultaneously applied. Thus, all photos are gradually arranged without empty space, and at the same time, their sizes almost become equal if there is not a change of conditions. Even if these two principles conflict, the algorithm will find a solution. Other principles related to geometric packing can be obtained similarly.

#### 4.2. Semantic mapping

As one of the simplest examples of semantic mapping,  $f_{attention}$  is a function to enlarge an interesting photo is represented by Eq. (9). Here, attention is a set of interesting photos given by adequate input device as a mouse or a gaze input device:

$$f_{attention}(I, \vec{P}) = A_{a1} \cdot \{Scale_{max} - Scale(P)\} \text{ if } P \in Attention \quad (9)$$

Eq. (10) shows how a focused photo attracts other ones with similar attributes. Here,  $Similarity(P_i, P_j)$  is the similarity between photos  $P_i$  and  $P_j$ , and if this value is larger than a threshold,  $P_j$  moves toward  $P_i$ , and away otherwise. The similarities are assumed to be calculated by feature values obtained by image processing or from tags of photos. Other related principles of semantic mapping can be obtained similarly.

$$f_{attraction}(I, \vec{P}) = \begin{cases} \sum_i^N A_{s2} \{Similarity(P, P_i) - Threshold\} \cdot \{Position(P_i) - Position(P)\} & \text{if } Similarity(P, P_i) \geq Threshold \\ \sum_i^N A_{s2} \{Similarity(P, P_i) - Threshold\} / \{Position(P_i) - Position(P)\} & \text{if } Similarity(P, P_i) < Threshold \end{cases} \quad (10)$$

Photographs can be sorted by the order of their attribute values based on Eq. (10), where  $O_{target}(P_{ij})$  is the correct order of  $P_i$  in terms of the  $j$ th attribute,  $O_{current}(P_{ij})$  is the current order of  $P_i$  in terms of the  $j$ th attribute, and  $Axis$  is the axis length with which the photographs are aligned.

$$f_{sort}(I, \vec{P}) = \{O_{target}(P_{ij}) - O_{current}(P_{ij})\} \cdot Axis \quad (11)$$

#### 4.3. Viewer's interactions

Even after the system reaches the balanced condition, the photograph behaviors can be observed when parameters of the display environment vary (e.g. when new photos are added or the size of the displaying window is changed). Also if a cursor (operated by mouse, gaze input, or other devices) is used, the photo overlaid by a cursor becomes larger using Eq. (9) with certain weight coefficients.

Users can observe the displayed photos on a computer display and interact with them at the same time by using cursors, touch panels, and gaze input devices. Multiple users may interact simultaneously with separate input devices. Keyword inputs, which are detected by a speech recognition technique, may also be an input method. On the other hand, the display resolution, the number of displays and their positions and orientations, also the size of the display are variables of the output. Such information about input and output is treated as  $I$  in previous equations.

### 5. Behaviours of photographs and performance

#### 5.1. Behaviors of photographs

The experimental system was developed in C#, as an integrated development environment, and Windows User API, Microsoft .NET Framework 3.5, DirectX SDK (November 2008), and Microsoft XNA Game Studio 3.1.

Some photograph behaviors using the principles explained in the previous section are illustrated here as well as to show the flexibility of our method. They are classified by types: geometric packing and semantic mapping.

### 5.2. Photograph behaviors with geometric packing

Fig. 2 shows how photos avoid overlaps when four equations (Eqs. (4)–(7)) are applied, and there are no interactions or other changes to environmental parameters. Fig. 2a shows the initial state where 75 photos are located randomly and overlapped. However, they gradually move to avoid overlaps and occupy the empty regions (using Eqs. (4), (5), as shown from left to right in Fig. 2. At the same time, the photos' sizes are varied (using Eqs. (6) and (7)), then soon become almost equal (Fig. 2d).

Fig. 3 is an example of photos avoiding overlaps by rotating in addition to the equations used in Fig. 2. In Fig. 3 (left) is the original layout, and on the right is the layout with collision-free arrangement with rotation. This is useful when photos are shown on tabletop surface displays shared by several users.

Fig. 4 shows an example of photograph behaviors when one of the environmental parameters, the window size, is changed. Once the window is enlarge, from Fig. 4a and b, the contained photos steadily move to the empty space (middle figure) according to Eq. (5) (as shown in Fig. 4c–e). This is gradually convergent with time so that the sizes of all photos become almost equal but avoid overlapping (Fig. 4f).

### 5.3. Photograph behaviors with semantic mapping

Fig. 5a shows an example where one photo is focused by overlaying a cursor (at the bottom center of the photo). Here the cursor is assumed to be controlled by an adequate input device. Fig. 5b shows photos arranged by color and user interests using the principles of geometric packing (i.e., Eqs. (4)–(7)). In this figure, there are two cursors (magenta and green) pointing at two photos (bottom-left daylight scene and upper-right night scene, respectively). Soon photos with similar colors are moved toward the focused ones. The final layout is achieved by Eq. (10) with semantic mapping principles as shown in Fig. 5b. Similarly, other feature values calculated by image processing can be used to a group of photos by applying this principle.

Fig. 6 shows an example of photos arranged using Geotags. In this example, photos arranged without overlapping (using the Eqs. (4)–(7)) are attracted by their geographical identification metadata (latitude and longitude coordinates) based on Eq. (10) and moved to their corresponding positions on a world map (as shown from (a) to (c) in Fig. 6).

Fig. 7 shows an example of finding someone's photos. Given a photo set of a social relationship, when a human face in a photo is selected, the size of that photo becomes larger by Eq. (9). Also, all photos containing the selected person are attracted and gathered around the focused photo dynamically by Eq. (10). Here, a face recognition function is assumed to be working and tags for the faces are adequately given in advance. Similarly, Fig. 8a displays examples of grouping photos with closed curves drawn by a mouse using meta-data given to each of the photos in advance. In this

figure, photos having meta-data of Mr. A and Mr. B are gathered in the red closed curve in the left and the green closed curve in the right, respectively. In the overlapping area of these two closed curves there are photos belonging to both Mr. A and Mr. B. Likewise, photographs having meta-data of animals and Australia are corrected in the red closed curve in the left and the green closed curve in the right, respectively, in Fig. 8b. In addition, photographs that have meta-data of both animals and Australia are corrected in the overlapping area of these two closed curves.

In addition, Fig. 9 shows an example in which about 480 photographs taken during almost 19 years of a family having two daughters are sorted from left to right in chronological order of the shooting date based on Eq. (11).

### 5.4. Discussions

A variety of photo arrangements can be achieved by replacing or adding functions without explicitly defining any prior relationships. As a result, even though the principles may conflict, our algorithm can always find a solution. In addition, compared with approaches based on combinatorial optimization, which always try to find a global optimal solution, our method generates solutions at every instant without fearing the local optimal solution. This is because we do not have to try finding the global optimal solution in the environment where conditions always dynamically vary.

In our method, each photo arrangement generated at every time instant may not always optimally satisfy all principles. However, these arrangements provide viewers with a dynamic photo-viewing environment where they can observe the smooth transitions of photo arrangements and the behaviors caused by their interactions.

Theoretically, the displayed photos by our algorithm are slightly vibrated because the function inherently includes a noise term in Eq. (1) that causes the system to be constantly in motion. In the examples described above, a fixed small amplitude was used for the noise. Larger amplitude noise increases the fluctuation of the environmental parameters, and such fluctuation is often useful for escaping local optimum. But if such large vibration is not suitable, the vibrations can be removed using a filter before rendering. In addition, the amplitude can be varied during photo arrangement to obtain a better performance.

### 5.5. Performance evaluation

In order to know the computational performance of the proposed algorithm, the frame rate of the implemented system was evaluated with various numbers of photos, and analyzed how multiple principles influenced the performance. Here, the experimental system was run on a Windows 7 64-bit PC with Intel Core i7 CPU (3.20 GHz) and 12.0 GB memory. It was not implemented for

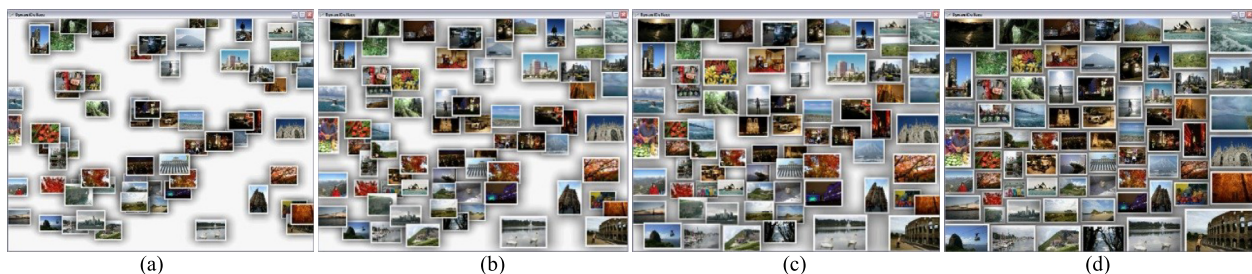


Fig. 2. Geometric packing of photos without the rotation.



Fig. 3. Geometric packing of photos with the rotation.

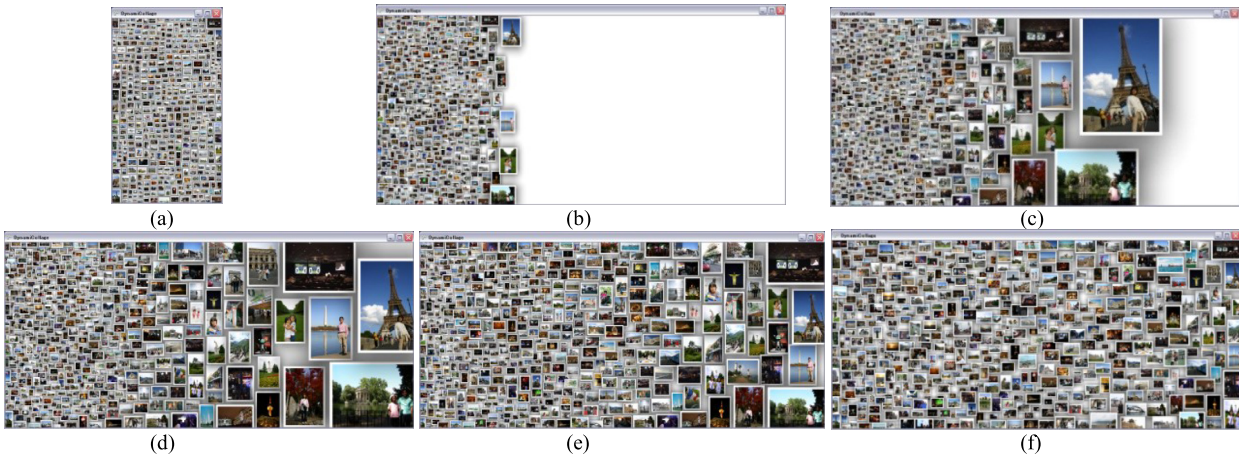


Fig. 4. Sequence when size of the displaying window is changed.



Fig. 5. Examples of geometric packing (a) and semantic mapping (b).

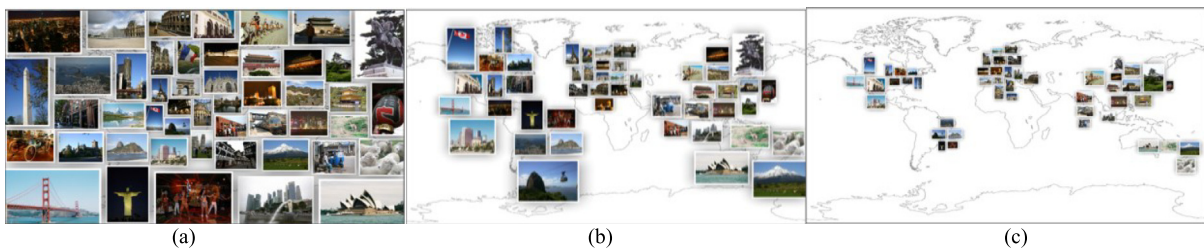


Fig. 6. Photographs arranged by Geotags.



**Fig. 7.** Finding someone's photographs. Given a photo set of social relationship and the human face in a photograph is selected by overlaying a pointer, all photographs containing the selected people are attracted and gather around the focused photograph dynamically.

parallel; therefore, it ran mainly on a single core of CPU. The graphics card used was NVIDIA GeForce GTX 285.

Regardless of the initial or convergent state, the computational complexity of the proposed algorithm is expected to depend on the number of photos and the principles of photo arrangement. Thus we computed the frame rates under various conditions (various numbers of photos and principles applied) and analyzed the factors that influence the performance of our method. In this experiment, standardized JPEG photos with  $512 \times 320$  pixels were used, and their display area on the screen was  $1024 \times 768$  pixels. Following four kinds of equations were denominationally applied: Eq. (4) for translating each photo to opposite direction when a collision occurs, Eq. (5) for molding each photo not to exceed the window's border, Eq. (6) for enlarging each photo as much as possible

without overlapping, and Eq. (7) for shrinking the larger one to avoid overlapping.

Average computation times (FPS, frame per second) for each ten-second processing cycles are plotted in Fig. 10. In this graph, the blue line represents result when all four principles are switched on. Magenta, cyan, and green lines represent results when one of the four principles is switched off one by one; the magenta line represents result with Eqs. (4)–(6), the cyan line represents result with Eqs. (5)–(7), the green line represents result with Eqs. 4, 5 and 7, and the red line represents result with only Eq. (5), respectively.

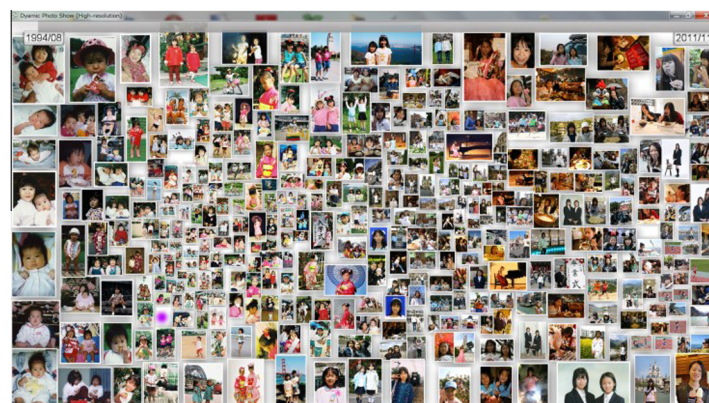
Note that the performance was saturated at around 60 FPS for a smaller number of photos (less than 600 or 700). However, when the number of photos exceeds 700, the performance drops rapidly. From these results, noticeable better performance can be obtained if Eq. (4) for translation and Eq. (6) for enlarging are not used because probabilities of photo overlap decrease. At the same time, the performance was deteriorated without using Eq. (7) for shrinking because the overlapping occurs more frequently. Therefore, we can conclude that the overlapping avoidance is crucial for the performance of our algorithm. In addition, even when only the equation for molding (Eq. (5)) is active, the frame rate is also not so high when the number of photos becomes larger (see red line in the graph), because the system need to determine intersections between each photo and the displaying window's border.

## 6. User evaluation of D-FLIP

Photo arrangements with dynamic motions, shown in previous section, are expected to be effective in many situations such as viewing many photos at once, surveying a set of photos from



**Fig. 8.** Examples of photographs corrected by meta-data.



**Fig. 9.** Example of photographs sorted in chronological order.

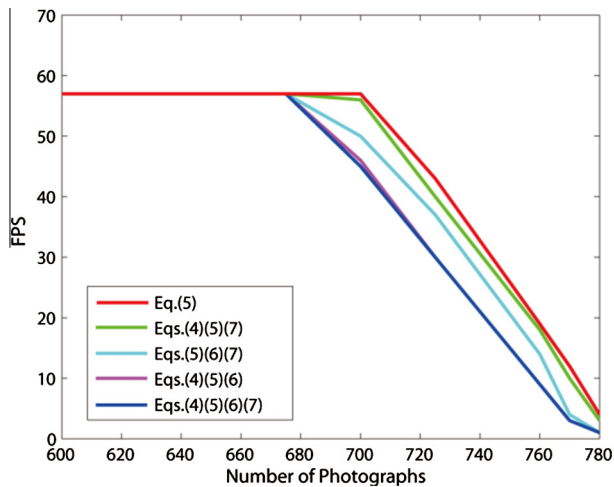


Fig. 10. Transition of frame rate.

different layouts (viewpoints), and finding pictures from the dynamic motions. Also, the smooth visual effect caused by interactions will keep up users' motivation to actively see and interact with photos. Our study investigates this further and explores whether it helps users perform browsing/sharing tasks easier.

We compared D-FLIP to the Windows Explorer (Explorer), a default Windows program. This is because although existing software (such as PhotoMesa [4], PhotoFinder, etc.) incorporate some of D-FLIP's animated properties, none of them supports animation of all photos in the collection. Instead, Explorer (of Windows 7) was chosen to be the most suitable candidate because: (1) it can be easily customized to have a separate area to store selected photos but still show the non-selected ones to support the story; (2) it includes many features of other photo browsing software such as: view all photos of the collection, sort or group photos (i.e. by name, tags, rating, etc.); and (3) it is a Windows built in software which is familiar and easy to use for users.

### 6.1. Task design

The experiment is a modified version of [43] where a user shares a story with a friend by showing her the collections.

Each experiment session required two participants: one narrator and one listener. The narrator brought two sets of photos which were taken by them from different events. The participants sat beside each other in front of a 27-inch monitor (2560 × 1440 resolution) displaying the narrator's personal photos (Fig. 11). The narrator then examined the photos and selected 10 which he thinks represent his/her story. Afterward, he narrates the story to



Fig. 11. Two participants in the D-FLIP mode.

the listener. The narrator was encouraged to use the selected photos to tell the story but can also use other non-selected photos to enrich it. The narrator was encouraged to tell about the actual events that happened in the photos as well as to interact with the program in a natural manner. During these steps, the listeners observed, listened, and enjoyed the story. We measured the task engagement from both narrator and listener to investigate the effect of the interactive application on the person who actually interacts with the program and on the person who does not interact but observe the interaction and listen to the story. Each participant wore an Emotiv EEG headset during the study.

There is a storyboard in D-FLIP to support the storytelling mode. The area has ten equal boxes which hold ten selected photos (Fig. 12, left). In case of Windows Explorer, we used two Windows Explorer windows with one window above another. The bottom window was used to store 200 photos and the top window was used to store the selected 10 photos (Fig. 12, right). In Windows 7, the layout of these windows was modified to display all photos in thumbnails and without filenames. All bars and additional panes (e.g. Library, Details, and Navigation) were hidden to make Windows Explorer comparable to D-FLIP in term of visualization and functions. Features of D-FLIP in this experiment included dynamic arrangement when overlaying a pointer; attraction for photos with similar colors; and timeline to display photos in chronological order (enabled manually by the narrator). These features were chosen as they are comparable with features of Windows Explorer. Moreover, the feature which enables similar color photos gathering was chosen because of the assumption that users usually take several photos of one scene in order to choose the best one afterward. Hence there are many collections of similar photos in each photoset.

### 6.2. Method

14 participants (5 females and 9 males) between the ages of 19 and 30 volunteered for the study and were arranged into 7 pairs. All were from the local university and had at least 5 years experience of using Windows Explorer for viewing photo collections. The narrators brought to the experiment their 400 photos divided into 2 sets of 200 each. These photos were resized to 640 × 480, rotated if necessary by the experimenter before beginning the study. Most narrators brought photos from their previous trips (e.g., a game exhibition or scenic photos of various places). All participants had normal or correct-to-normal vision. The entire experiment took about 1 h 15 min and participants were paid for their time (about 10 USD).

First, the narrator had adequate time to practice with D-FLIP and Windows Explorer using a sample photoset (Fig. 12). After this they wore the Emotiv headsets. Each experiment session had two blocks each with either D-FLIP or Windows Explorer. Each block started with a 3-minute reference recording where the participants then sat still, comfortably, and with eyes opened. After this, the narrator prepared a story in 5 min by selecting 10 photos from her first photoset. The narrator then told the story to the listener within 5 min. Both participants had a 2-minute break before the same procedure was repeated with the next program (as demonstrated in Fig. 13, left). NASA-TLX forms were given to both participants after each block. These forms collected ratings of mental demand, physical demand, performance, effort, and frustration. The order of the two programs was balanced between experiment sessions.

In the storytelling step, we only recorded EEG signals from the listener. This was because the narrator needed to speak freely with facial, hand, and body movements which would contaminate the EEG signals. To reduce the noise in EEG data (we also decontaminate the signal) we asked: (1) the narrator to not wait for feedback



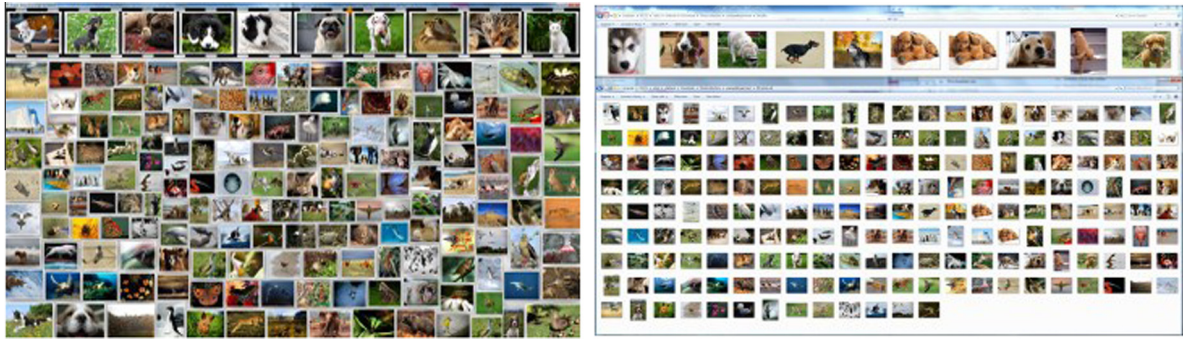


Fig. 12. Storytelling mode with D-FLIP (left) and Windows Explorer (right).

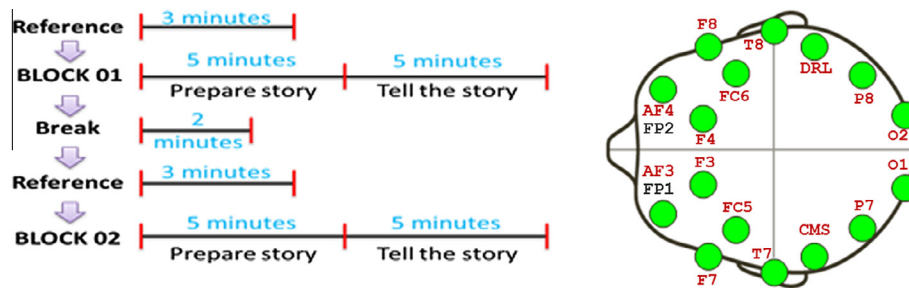


Fig. 13. Experiment procedure (left) and Emotiv channels in the 10–20 international system (right).

from the listener; (2) the listener to minimize body or facial movements as feedbacks to the story; and (3) both participants to focus on the monitor. The narrator was encouraged to use both programs in a natural manner.

### 6.3. Data acquisition and analysis

We used an Emotiv EPOC wireless headset for measuring EEG signals. This headset can capture EEG signals at 14 channels and has 2 reference channels (DRL and CMS) as shown in Fig. 13, right. Its signals quality have been validated by previous studies (e.g. detecting and classifying Event Related Potentials [29,44,45], evaluating visualization effectiveness [46], exploring the nature of decision making [7], and designing a BCI controlled video game [47]). This headset is portable, easy to setup, and most users can wear it comfortably for at least an hour [48].

EEG signals were sampled at 2 KHz internally then down-sampled to 128 Hz output. We inserted three types of markers into the marker channel, which is in parallel with EEG signal channels. These were: start recording, stop recording, and eye blinking events. We then adapted the “sliding window technique” which is commonly used to process EEG signals (e.g. [49,50]). The recorded EEG signals were segmented into 4-second epochs with 2-second overlaps between them. The DC offset was removed by subtracting the averaged value from the reference signals.

#### 6.3.1. EEG signals decontamination

To remove artifacts from muscle movements and eye blinks, we adapted a decontamination process [35] to decontaminate Emotiv’s EEG signals. These steps are: (1) Three data point spikes detection with amplitudes greater than 40  $\mu$ V (e.g. caused by tapping or bumping the sensors); (2) Notch, low and high pass filters to remove unwanted artifacts (e.g. from muscle movements); (3) Detect and remove eyes blinks. To aid in step (3), we used Emotiv SDK to insert markers when a user blinked. Two independent marker channels for the narrator and listener were used to remove eye-blinks [35].

We then performed a validation step where epochs were removed if the combined power of the *alpha*, *beta*, or *theta* frequency bands of an epoch changed more than 20% of its original values. We removed totally 7.83% of collected signals in which 5.13% due to excessive signals and 2.7% due to high changes in combined power spectral.

#### 6.3.2. Task engagement calculation

Engagement index was calculated as in Pope et al. [38].  $\beta$ ,  $\alpha$ , and  $\theta$ , are the combined power in the ranges of 13–22 Hz, 8–12 Hz and 5–7 Hz frequency bands.

$$\text{Engagement Index} = \beta / (\alpha + \theta) \quad (12)$$

We used F3, F4, FC5, FC6, P7, P8, O1, and O2 channels as they are the surrounding channels of Cz, Pz, P3, and P4 which were used by Pope et al. [38]. Engagement indexes were calculated in 4s window for each task & user separately and then normalized for the narrator or listener.

#### 6.3.3. Results and discussion

Table 1 summarizes the results, which are the averaged task engagement during the session, for two types of tasks (story preparation and story telling), two types of programs (Windows Explorer and D-FLIP), and with two participant types (Narrator and Listener).

Fig. 14 (left) shows a time series sample of engagement during the preparation task of one narrator. Fig. 14 (right) shows the task

**Table 1**  
Average task engagement for narrators and listeners.

Tasks	Programs	Narrator	Listener
Story preparation	Windows Explorer	0.603	0.480
	D-FLIP	0.554	0.436
Story telling	Windows Explorer		0.512
	D-FLIP		0.440

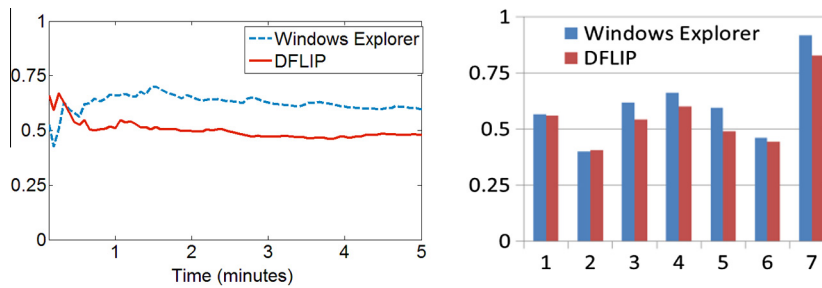


Fig. 14. An example of a narrator's task engagement changes in time in one session (left); and task engagement of each narrator (right).

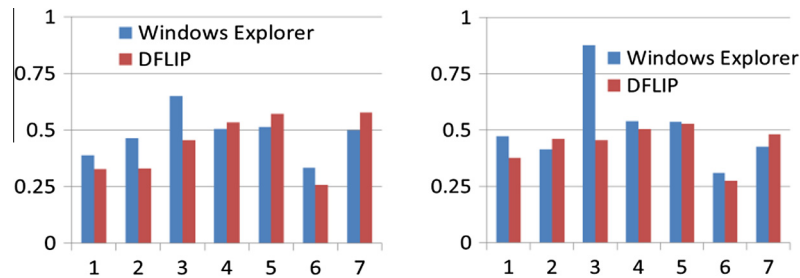


Fig. 15. Task engagements of listeners in the story preparation task (left) and story telling task (d) between Windows Explorer and D-FLIP.

engagement values of narrators in each experiment session between two programs. *T*-test shows that there is a significant difference between those two ( $t = 2.95, p < 0.05$ ). However, we found no significant differences between two types of program in listeners ( $p > 0.5$ ). The details are shown in Fig. 15 (left – story preparation) and Fig. 15 (right – story telling). This may be due to the individual differences where the listeners did not directly interact with the programs hence the measured task engagement mainly depended on the storytelling skills of the narrators.

From the NASA-TLX, we found a significant difference in mental demand of the narrator between two tasks: Windows Explorer (mean: 9.14) and D-FLIP (mean: 5.00). However, we found no significant difference in physical demand, temporal demand, performance, effort, and frustration of narrator between two tasks ( $p > 0.5$ ). For the listener, no significant differences were found in mental demand, temporal demand, effort, and frustration between two tasks.

The results showed that the narrators had higher engagement when using the Windows Explorer compared to D-FLIP. This also means that Windows Explorer requires higher task loads hence users need to put more efforts to complete the same task in the same amount of time compared with D-FLIP. Consequently, besides the benefit of having more interest, D-FLIP can help users perform the task easier compared to a common program such as Windows Explorer. However, we only found the significant difference in the narrator's task engagement, not the listener's. Obtained results from the NASA-TLX also showed consistent results compared to the captured EEG signals. A possible reason is because only the narrators were the ones actually interact with the programs while the listeners were not. Consequently, with less task engagement or workload required in performing the task, users will find out that interacting easier and more pleasure.

In addition, although adequate practice time was given to the narrators prior to performing the task, there was possibility that the users felt more familiar with Windows Explorer and performed the task more skillful than D-FLIP. As a result, the actual difference between measured task engagements of two programs might be

larger if another program was used instead of Windows Explorer (i.e. a less famous program such as PhotoMesa). Our results show that even if there are effects of distraction and familiarity, D-FLIP still requires less task engagement hence less task loads and mental demands.

## 7. Discussion

Our user evaluation points out that D-FLIP requires lower level of mental effort and task engagement compared to Windows Explorer. Several factors of D-FLIP may contribute to this lower produced task engagement such as: high visibility with variety of layout, dynamic and smooth motions of photos, gathering photos based on similar attributes. As a result, D-FLIP requires less effort from users to perform the task hence they can focus on the story contents and interactions with other users.

It can be said that lower investment of efforts and task engagement in an interactive task also means that the task is easier to perform. However, if these values are too low users may feel bored and lose interest. Equivalently if the values are too high it can cause anxiety [51]. Our results show that D-FLIP helps users perform the task easier than a common program such as Windows Explorer. However, while taking advantage of this easiness, D-FLIP still keeps motivation and interest in users due to the dynamic, flexible, and interactive motions of photos produced by the proposed principles.

This combined with the fact that animated graphics enhances emotional connection with the system, makes D-FLIP a particularly powerful system for visualizing large collections of images.

Through performance evaluation, our proposed algorithm can work well with at least 700 photos. However, this can be easily improved further with a better computing hardware and parallelization for optimized implementation.

The smooth motion of each photo is characterized by (1) the implemented photo arrangement principles and (2) parameters values used in the equations (Eqs. (1)–(11)). These are implemented by careful considerations and preliminary experiments.

Our next step is to improve the algorithm categorizing the parameters and optimizing it for different types of interactive applications. Doing this will let the algorithm easily and flexibly adapt to various types of content. Besides existing content types (i.e. text, broadcasts, movies, web content, music), our algorithm can adapt to work with new and emerging types of content such as dreams visualization [52] where values of parameters are from fMRI patterns for an image. The images depicting a person's dream will have smooth and dynamic motions that make our visualization well suited for visualizing this type of data. This can also help to build an ecosystem of photos that includes various promising features such as pigeonholing, printing, automatic acquisition of meta-data, evolving into photo sharing sites, and coordinating with social network services. We can also explore other interaction devices such as digital tables with multi-touch interaction, voice recognition, brain-machine interface, range image sensors, and other sensing devices to further enhance the fluidity of interaction in different application contexts.

Although the focus of this paper is not on a novel evaluation methodology, we believe that our way of measuring task engagement using EEG offers greater insights into workings of an application. Our measured task engagement is consistent with the NASA-TLX results; providing a source of external validity to our measurement mechanism. This evaluation method can be used in other settings for similar comparisons. It can also be improved to capture emotions (e.g. relaxation, meditation) and users' inner states (e.g. error awareness).

## 8. Conclusion

We presented D-FLIP a program that displays a set of photos flexibly, dynamically, and interactively. The underlying algorithm adjusts the displaying principles adaptively when users interact with the program. Our performance evaluation shows it can handle smoothly at least 700 photos. Our user evaluation shows that D-FLIP requires less task engagement and mental effort from users allowing them to enjoy the content rather than manage it.

Our future works include meta-data auto acquisition, photo sharing site development, and developing an ecosystem of photographs and applications (e.g. recommender systems, street digital signage, etc.). Moreover, our proposed algorithm can be written in any programming language including web programming (e.g. JavaScript). This opens a wide range of web applications with various and rich content types.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.entcom.2014.08.005>.

## References

- [1] D. Kirk, A. Sellen, C. Rother, K. Wood, Understanding photowork, in: CHI, ACM, Canada, 2006, pp. 761–770.
- [2] C. Shen, N. Lesh, F. Vernier, Personal digital historian: story sharing around the table, *Interactions* 10 (2003) 15–22.
- [3] M. Balabanović, L.L. Chu, G.J. Wolff, Storytelling with digital photographs, in: CHI, ACM, The Hague, The Netherlands, 2000, pp. 564–571.
- [4] B.B. Bederson, PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps, in: UIST, ACM, USA, 2001, pp. 71–80.
- [5] R. Dachsel, M. Frisch, M. Weiland, FacetZoom: a continuous multi-scale widget for navigating hierarchical metadata, in: CHI, ACM, Italy, 2008, pp. 1353–1356.
- [6] J. Kustanowitz, B. Shneiderman, Meaningful presentations of photo libraries: rationale and applications of bi-level radial quantum layouts, in: JCDL, ACM, USA, 2005, pp. 188–196.
- [7] A. Graham, H. Garcia-Molina, A. Paepcke, T. Winograd, Time as essence for photo browsing through personal digital libraries, in: JCDL, ACM, USA, 2002, pp. 326–335.
- [8] S. Ahern, M. Naaman, R. Nair, J.H.-I. Yang, World explorer: visualizing aggregate data from unstructured text in geo-referenced collections, in: JCDL, ACM, Canada, 2007, pp. 1–10.
- [9] A. Gomi, T. Itoh, A personal photograph browser for life log analysis based on location, time, and person, in: SAC, ACM, Taiwan, 2011, pp. 1245–1251.
- [10] P. Brivio, M. Tarini, P. Cignoni, Browsing large image datasets through voronoi diagrams, *IEEE Trans. Visual. Comput. Graphics* 16 (2010) 1261–1270.
- [11] T. Itoh, Y. Yamaguchi, Y. Ikehata, Y. Kajinaga, Hierarchical data visualization using a fast rectangle-packing algorithm, *IEEE Trans. Visual. Comput. Graphics* 10 (2004) 302–313.
- [12] A. Gomi, R. Miyazaki, T. Itoh, L. Jia, CAT: A Hierarchical image browser using a rectangle packing technique, in: Information Visualisation, 2008. IV '08. 12th International Conference, 9–11 July 2008, pp. 82–87.
- [13] J.-C. Chen, W.-T. Chu, J.-H. Kuo, C.-Y. Weng, J.-L. Wu, Tiling slideshow, in: MM, ACM, CA, USA, 2006, pp. 25–34.
- [14] C. Rother, L. Bordeaux, Y. Hamadi, A. Blake, AutoCollage, in: SIGGRAPH, ACM, USA, 2006, pp. 847–852.
- [15] J. Wang, L. Quan, J. Sun, X. Tang, H.-Y. Shum, Picture Collage, in: CVPR, IEEE Computer Society, 2006, pp. 347–354.
- [16] C. Rother, S. Kumar, V. Kolmogorov, A. Blake, Digital tapestry, in: IEEE Computer Society Conference on CVPR, IEEE Computer Society, 2005, pp. 589–596.
- [17] C.W. Reynolds, Flocks, herds and schools: A distributed behavioral model, *SIGGRAPH Comput. Graphics* 21 (1987) 25–34.
- [18] D. Helbing, P. Molnár, Social force model for pedestrian dynamics, *Phys. Rev. E* 51 (1995) 4282–4286.
- [19] A.V. Moore, Time-varying data visualization using information flocking boids, in: INFOVIS, 2004, pp. 97–104.
- [20] G. Turk, Generating textures on arbitrary surfaces using reaction-diffusion, *SIGGRAPH Comput. Graphics* 25 (1991) 289–298.
- [21] K. Shimada, D.C. Gossard, Bubble mesh: automated triangular meshing of non-manifold geometry by sphere packing, in: 3SOLID95, ACM, Salt Lake City, Utah, USA, 1995, pp. 409–419.
- [22] T.M.J. Fruchterman, E.M. Reingold, Graph drawing by force-directed placement, *Softw. Pract. Exp.* 21 (1991) 1129–1164.
- [23] J. Heer, S.K. Card, J.A. Landay, Prefuse: a toolkit for interactive information visualization, in: CHI, ACM, Portland, Oregon, USA, 2005, pp. 421–430.
- [24] W. Li, P. Eades, N. Nikolov, Using spring algorithms to remove node overlapping, in: APVIS, Australian Computer Society Inc, Sydney, Australia, 2005, pp. 131–140.
- [25] S. Kim, M. Yoon, S.M. Whang, B. Tversky, J.B. Morrison, The effect of animation on comprehension and interest, *JCAL* 23 (2007) 260–270.
- [26] L.P. Rieber, Animation, incidental learning, and continuing motivation, *J. Educat. Psychol.* 83 (1991) 10.
- [27] W. Kintsch, Learning from text, levels of comprehension, or: why anyone would read a story anyway, *Poetics* 9 (1980) 87–98.
- [28] D.-S. Ryu, W.-K. Chung, H.-G. Cho, PHOTOLAND: a new image layout system using spatio-temporal information in digital photos, in: SAC, ACM, Sierre, Switzerland, 2010, pp. 1884–1891.
- [29] C. Vi, S. Subramanian, Detecting error-related negativity for interaction design, in: CHI, ACM, Austin, Texas, USA, 2012, pp. 493–502.
- [30] S.I. Hjelm, Research + design: the making of Brainball, *Interactions* (2003) 26–34.
- [31] D. Hagemann, J. Hewig, C. Walter, A. Schankin, D. Danner, E. Naumann, Positive evidence for Eysenck's arousal hypothesis: a combined EEG and MRI study with multiple measurement occasions, *J. Personal. Individ. Differ.* (2009) 5.
- [32] C. Berka, D.J. Levendowski, M.N. Lumicao, A. Yau, G. Davis, V.T. Zivkovic, R.E. Olmstead, P.D. Tremoulet, P.L. Craven, EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks, *ACEM* 78 (2007).
- [33] G. Matthews, S.E. Campbell, S. Falconer, L.A. Joyner, J. Huggins, K. Gilliland, R. Grier, J.S. Warm, Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry, *Emotion* 2 (2002) 315–340.
- [34] G. Matthews, J. Warm, L. Reinerman, L. Langheim, D. Saxby, Task engagement, attention, and executive control, in: A. Gruszka, G. Matthews, B. Szymura (Eds.), *Handbook of Individual Differences in Cognition*, Springer, New York, 2010, pp. 205–230.
- [35] C. Berka, D.J. Levendowski, R.E. Olmstead, M.V. Popovic, M. Cvetinovic, M.M. Petrovic, G. Davis, M.N. Lumicao, P. Westbrook, Real-time analysis of EEG indices of alertness, cognition, and memory with a wireless EEG headset, *IJHCI* 17 (2004) 151–170.
- [36] M. St John, D.A. Kobus, J.G. Morrison, DARPA augmented cognition technical integration experiment (TIE), in: US Navy SPAWAR Systems Center, San Diego, 2003.
- [37] C. Berka, D.J. Levendowski, C.K. Ramsey, G. Davis, M.N. Lumicao, K. Stanney, L. Reeves, S.H. Regli, P.D. Tremoulet, K. Stibler, Evaluation of an EEG-workload model in an Aegis simulation environment, in: SPIE Defense and Security Symposium, Biomonitoring for Physiological and Cognitive Performance during Military Operations, 2005, pp. 90–99.
- [38] A.T. Pope, E.H. Bogart, D.S. Bartolome, Biocybernetic system evaluates indices of operator engagement in automated task, *Biol. Psychol.* 40 (1995) 187–195.
- [39] F.G. Freeman, P.J. Mikulka, L.J. Prinzel, M.W. Scerbo, Evaluation of an adaptive automation system using three EEG indices with a visual tracking task, *Biol. Psychol.* 50 (1999) 61–76.

- [40] L.J. Prinzel, F.G. Freeman, M.W. Scerbo, P.J. Mikulka, A.T. Pope, A closed-loop system for examining psychological measures for adaptive task allocation, *Int. J. Aviat. Psychol.* 10 (2000) 393.
- [41] F. Freeman, P. Mikulka, M. Scerbo, L. Prinzel, K. Cloutre, Evaluation of a psychophysiological controlled adaptive automation system, using performance on a tracking task, *Appl. Psychophysiol. Biofeedback* 25 (2000) 103–115.
- [42] P.J. Mikulka, M.W. Scerbo, F.G. Freeman, Effects of a biocybernetic systems on vigilance performance, *Hum. Factors* 44 (2002) 654.
- [43] K. Rodden, K.R. Wood, How do people manage their digital photographs?, in: CHI, ACM, Ft. Lauderdale, Florida, USA, 2003, pp. 409–416.
- [44] A. Campbell, T. Choudhury, S. Hu, H. Lu, M.K. Mukerjee, M. Rabbi, R.D.S. Raizada, NeuroPhone: brain-mobile phone interface using a wireless EEG headset, in: ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds, 2010, pp. 3–8.
- [45] C.T. Vi, I. Jamil, D. Coyle, S. Subramanian, Error related negativity in observing interactive tasks, in: CHI, ACM, Toronto, Canada, 2014.
- [46] E.W. Anderson, K.C. Potter, L.E. Matzen, J.F. Shepherd, G.A. Preston, C.T. Silva, A user study of visualization effectiveness using EEG and cognitive load, in: EuroVis, Eurographics Association, Bergen, Norway, 2011, pp. 791–800.
- [47] M.G. Ames, L. Manguy, PhotoArcs: ludic tools for sharing photographs, in: MM, ACM, Santa Barbara, CA, USA, 2006, pp. 615–618.
- [48] J.I. Ekanem, T.A. Davis, I. Alvarez, M.T. James, J.E. Gilbert, Evaluating the ergonomics of BCI devices for research and experimentation, *Ergonomics* 55 (2012) 592–598.
- [49] J.C. Lee, D.S. Tan, Using a low-cost electroencephalograph for task classification in HCI research, in: UIST, 2006, pp. 81–90.
- [50] C.W. Anderson, Z. Sijerčić, Classification of EEG signals from four subjects during five mental tasks, in: EANN, Springer, 1996, pp. 407–414.
- [51] G. Chanel, C. Rebetez, M. Bétrancourt, T. Pun, Boredom, engagement and anxiety as indicators for adaptation to difficulty in games, in: Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era, ACM, Tampere, Finland, 2008, pp. 13–17.
- [52] T. Horikawa, M. Tamaki, Y. Miyawaki, Y. Kamitani, Neural decoding of visual imagery during sleep, *Science* (2013).