Entertainment Computing 5 (2014) 451–462

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: ees.elsevier.com/entcom

A dynamic flexible and interactive display method of digital photographs ☆

Chi Thanh Vi [a,*], Kazuki Takashima [b], Hitomi Yokoyama [b,1], Gengdai Liu [c], Yuichi Itoh [d], Sriram Subramanian [a], Yoshifumi Kitamura [b]

[a] Department of Computer Science, University of Bristol, Woodland Road, Bristol BS8 1UB, United Kingdom
[b] Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
[c] OLM Digital Inc., Mikami Bldg, 2F, 1-18-10 Wakabayashi, Setagaya-ku, Tokyo 154-0023, Japan
[d] Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

A R T I C L E I N F O

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)

A B S T R A C T

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.

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.

☆ This paper has been recommended for acceptance by Haruhiro Katayose.

* Corresponding author.
E-mail addresses: Chi.Vi@bristol.ac.uk (C.T. Vi), takashima@riec.tohoku.ac.jp (K. Takashima), hitomi-y@cc.tuat.ac.jp (H. Yokoyama), liugengdai@gmail.com (G. Liu), itoh@ist.osaka-u.ac.jp (Y. Itoh), Sriram.Subramanian@bristol.ac.uk (S. Subramanian), kitamura@riec.tohoku.ac.jp (Y. Kitamura).
[1] Tokyo University of Agriculture and Technology, Graduate School of Engineering, Advanced Information Sciences Division, 184-8588 Koganei-shi, Naka, Tokyo 2-24-16, Japan.

http://dx.doi.org/10.1016/j.entcom.2014.08.005
1875-9521/

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' metal 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]).