Visual Informatics 1 (2017) 16–24

Contents lists available at ScienceDirect

Visual Informatics

journal homepage: www.elsevier.com/locate/visinf

Image grid display: A study on automatic scrolling presentation

Marco Porta *, Stefania Ricotti

Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy

ARTICLE INFO

Article history:
Available online 18 January 2017

Keywords:
Image grid display
Automatic scrolling
Image presentation modes
Image collections
Rapid serial visual presentation

ABSTRACT

In this paper we describe a study on image grid display with automatic vertical scrolling. While scroll operations are normally carried out manually by the user, in the context of RSVP (Rapid Serial Visual Presentation) techniques this work considers a presentation mode in which the image grid is automatically scrolled. Through experiments carried out with 50 testers, we have investigated user performance while looking for specific target subjects within large collections of images. Different numbers of columns and scrolling speeds have been considered. The search task implied both clicking on the identified target pictures and simply vocally stating their visual recognition. To this purpose, and to identify possible specific gaze behaviours, eye tracking technology has been exploited. The obtained results show that number of columns and scroll speed do affect search performance. Moreover, the user's gaze tends to focus on different screen areas depending on the values of these two parameters. Although it is not possible to definitely find an optimal columns–speed combination that is valid in all cases, the particular context of use can suggest feasible solutions according to one's needs. To the best of our knowledge, image grid display with automatic scrolling has never been studied to date.

1. Introduction

Grid (or tabular) display is certainly the most common way of presenting groups of images on a screen. For instance, so-called "thumbnail" pictures arranged in rows and columns can be found in online previews (e.g. result pages of 'Google image search' or photos in Instagram profiles), folder content display (e.g. MS Windows medium, large, and extra-large icons), or graphic tools (e.g. in the 'Styles' palette of Adobe Photoshop). The main advantage of the grid is evident: since images are normally rectangular, it allows an efficient exploitation of the available space. It is also intuitive, as the table layout is used to display many kinds of data, outside the context of pictures as well.

Image grids come in a variety of configurations, in which image size, number of columns and number of rows are the most common variables. Of course, the smaller the pictures the higher their display density on the screen. However, small graphic representations may be difficult to identify, especially when the recognition of specific images or subjects is required. A trade-off is therefore necessary between the number of presented images and their size, which usually depend on factors such as presentation goals, kind of subjects, and needed search accuracy.

Unless the number of pictures is very low, vertical scroll operations are almost always necessary. Alternatively, different sets of images can be displayed in separate "screens", as sometimes happens in websites. However, especially online, loading new screens may result in a slower image browsing experience, which may not be always acceptable. A combination of the two solutions—different screens containing scrollable grids—is an adequate choice in many cases.

In this paper, we consider an unusual approach to image grid display with vertical scrolling. While scroll operations are generally performed manually by the user, here we focus on a presentation mode in which the grid is *automatically* scrolled up with a constant speed. In other words, the user does not need to employ the mouse or the keyboard to move the displayed set of images, which is automatically shifted upwards. The rationale for this is twofold.

Firstly, an automatic scroll can be seen as a simulation of a manual scroll with constant scroll rate. Thus, indirectly, it is possible to study user behaviour while performing one of the most common—albeit typically "unconscious"—operations in grid display with many images. For experimental purposes, using the same speed with all testers can provide more reliable results (at least for the goals of our investigation).

The second and most important reason for which we decided to study scrolling image grids relates to *Rapid Serial Visual Presentation* (RSVP) modes (Spence, 2002). Putting it simple, RSVP

* Corresponding author.
E-mail address: marco.porta@unipv.it (M. Porta).
Peer review under responsibility of Zhejiang University and Zhejiang University Press.

http://dx.doi.org/10.1016/j.visinf.2017.01.003

means displaying many images in rapid succession on the screen. The purpose of this visualization strategy is to present numerous pictures in a short time, so that users can rapidly find what they are looking for. While in *indexed* image databases pictures have textual information associated with them, and in *Content-Based Information Retrieval* (CBIR) systems pictures are identified based on intrinsic characteristics such as colour, shape and texture, in many cases the manual browsing of very large collections of images is necessary. In fact, there are several situations in which neither indexed nor CBIR systems are available, or we need to select only some pictures because they are characterized by specific features or simply because we "like" them (e.g. for decorative purposes). Browsing big collections of images with traditional display modes—e.g. the static grid—may result in a very boring task, and RSVP can be the right solution.

Several variants of the basic RSVP method (in which single pictures, one at a time, are rapidly displayed on the screen) have been proposed (Cooper et al., 2006; Porta, 2006, 2009). For example, in the *mixed* presentation mode images are displayed in groups of four or more; in the *diagonal* mode, images move from a corner of the screen to the opposite one; in the *collage* mode, pictures appear in random positions of the screen, like if thrown on a table; in the *volcano* mode, images emerge from the Centre of the screen and move radially towards the edges; in the *fountain* mode, pictures are randomly "spurted" upwards and then fall back.

Although some general guidelines have been drawn from experimental evidences (Witkowski and Spence, 2012; Spence and Witkowski, 2013), each RSVP solution is characterized by its own features and needs customized investigations for its potentials and drawbacks to be really understood.

To the best of our knowledge, image grid display with automatic scrolling has never been studied to date. In this paper, we therefore focus on this presentation mode and investigate its possible use as an RSVP method. In particular, we consider two main variables that affect its performance, namely *number of columns* and *scroll speed*. While also image size may be an important factor, in our experiments we opted for a constant value, so as to limit the complexity of the analysis. Observing the heights of pictures displayed, for example, in the result pages of the main search engines (which range from about 140 pixels of Bing to about 170 of Yahoo! and 180 of Google), in MS Windows large icon display (90), or in YouTube home page video thumbnails (110, 170 including the description), we can notice a relatively high variability. Even if our choice – 150 × 150 pixels—cannot account for all possible dimensions, we think it is a good compromise between widespread sizes and acceptable accuracy in the recognition of image subjects.

For our analysis, we also exploited *eye tracking* technology (Duchowski, 2007). An eye tracker is a device capable of recording one's gaze direction while looking at a screen, thus allowing precise understanding of what the user is watching during experimental tests. Eye movements occur as very fast *saccades* followed by *fixation* periods of about 100–600 ms, during which the eye is almost still. Eye tracking data have provided us with interesting insights into gaze behaviours and their relationships with the considered variables.

In summary, this paper tries to answer two main research questions regarding image grid display with automatic scrolling:

- Q1 Is there any best combination of number of columns and scroll speed that guarantees a "good" performance when the purpose is to find specific subjects within an image set?
- Q2 Is image search characterized by any particular gaze behaviours? And, if so, do these behaviours depend on number of columns and scroll speed?

The article is structured as follows. In Section 2 we describe some previous works directly or indirectly connected with grid layout image display and RSVP presentation modes. In Section 3 we then illustrate our study, in terms of participants and task description. Subsequently, in Section 4, the obtained results are presented in relation to user performance, gaze behaviour and gaze scan path length. Lastly, in Section 5, we discuss the achieved outcomes and draw some conclusions, also providing hints for future research.

2. Related work

Although there seems not to be any study specifically devoted to scrolling image grid display—let alone investigations focused on automated scroll—there are works focused on variants of the grid display, rapid serial visual presentations, and the use of eye tracking as an evaluation tool or an active control mechanism for information visualization: here we provide some representative examples of them.

PhotoFinder (Shneiderman and Kang, 2000) and PhotoMesa (Bederson, 2001) are two cases of early and well-known visualization techniques based on the grid. The first is a photo annotation tool that allows the user to add captions and edit images, while the second is an application in which multiple directories of images can be viewed within a zoomable environment. Solutions have also been proposed (Igarashi and Hinckley, 2000) which dynamically control grid zooming depending on the scrollbar speed (the faster the scrollbar, the less the zoom applied): since the grid display usually requires scrolling, small movements of the scrollbar may produce big shifts of the grid if it contains many images.

Often, pictures are also clustered according to some criteria in more or less standard grid arrangements, so that they can be hierarchically browsed in non-linear manners or arranged according to their mutual similarity. For instance, Liu et al. (2004) analyse user needs for web image search and propose a similarity-based organization to present search results. Ren et al. (2009) describe an interactive interface in which images are clustered through an unsupervised graph-based clustering algorithm. Strong et al. (2010) present an approach supporting dynamic zooming in which visually similar images are displayed in the same locations (either scattered or aligned to a grid, depending on the selected display mode). Kleiman et al. (2015) describe a system in which images are dynamically arranged (on-the-fly, depending on users' navigation tendencies and interests) close to their nearest neighbours in a high-dimensional feature space.

A variant of the pure grid arrangement is presented in a work by Schaefer (2010), where a similarity-based picture organization approach is used to display images onto a sphere, with which the user can interact. Likewise, a 3D interface is used by Schoeffmann and Ahlström (2012) to display image thumbnails on a cylinder, based on visual similarity.

As regards RSVP approaches (some of which connected with eye tracking), Fan et al. (2003) developed a prototype solution for browsing large amounts of images which exploited a gaze-based attention model. Oyekoya and Stentiford (2006) studied a gaze driven image retrieval system and tested it with 13 users who had to find target images within 4 × 4 screens, with the screen automatically changing when the duration of all fixations on a picture reached a threshold. Results indicate a slower mouse response compared to the eye tracking approach. Corsato et al. (2008) compared four RSVP techniques (Floating, Collage, Volcano, and Shot) in the specific task of finding the highest number of pictures matching a textual description, similarly to our experiments. The 30 testers involved were studied through an eye tracking system with the main aim to validate the viability of a gaze controlled image selection method. Besides ordinary pictures, other studies