Abstract—Digital cameras are replacing analogue ones, and electronic photo albums are everywhere on the Internet. Image browsing is used for searching the contents of such albums, but also for a variety of other tasks.

Results from image browser studies are sometimes either unexpected or in conflict with previous results. Based on observations made during such experiments, we have identified and analysed a number of visual factors that seem to be relevant for the outcome of the test. This analysis identifies research gaps that need to be filled in order to support choice of appropriate browsing technique in interface design. Some suggestions for future user tests are also given.

Index terms—Fisheye, Fitts’ law, icons, image browsing, usability, zoom.

I. INTRODUCTION

Most computer users would agree that the amount of screen space can be insufficient, but never too large. One type of task that requires a lot of screen space is viewing images (Figure 1). As digital cameras are becoming more common, people tend to store their images in electronic albums, rather than in physical ones. People also more often use the Internet in search for information about geographic locations (Figure 2). Normally some browsing technique is used, that enlarges images and enables overviews of entire sets of images as well as selected images in detail.

Browsing techniques have been compared in several studies over the years, with various results. The question of which technique is most efficient gets an answer in each of these studies, but the answers differ from one study to another. The analyses do not really form a clear pattern that can provide guidelines for future design. Before conducting yet another test, we wanted to look into the details behind what makes a certain technique efficient.

Image browsing normally means either or both of the following:

- A set of images is displayed on the screen, and the task is to identify a possible target image, usually by enlarging it with some technique.
- An image, too large to fit the screen at readable size, is displayed, and the task is to investigate it.

Figure 1 Electronic photo album
Distortion-oriented techniques are interfaces that provide both focus and context, such as the perspective wall [16], the bifocal view, and all types of fisheye views [8], [15]. Focus is normally moved either by dragging or clicking.

Multiple views are often used in combination with another technique. Especially a zoom interface can benefit from an overview to prevent the user from getting lost.

There are mainly two ways to improve the efficiency and usability of image browsing: Modify the presentation of the data, or modify, or change, the browsing technique used for the task.

It seems as if certain image browsing techniques benefits from certain layout and presentation strategies. This paper aims at identifying and understanding factors that affect usability for existing techniques. Perhaps this analysis could help interpreting the outcome of image browser tests, and give some useful directions on how to conduct future studies with more homogeneous results. First we will give a summary of comparisons made, followed by a definition of task completion time for image browsing present. After that we analyse the visual factors of image browsing, and discuss the impact of each of them. Finally, we suggest some topics for future work.

II. RELATED WORK

Several studies comparing browsing techniques have been conducted, with various results, not always supporting previous work.

In a comparison by Beard and Walker [1] interfaces using scrollbars, zoom, and roam (pan) with and without overview, were compared. They found that all techniques gave better results with the overview, and that scroll bars were significantly slower than the other two techniques.

A study by Kaptelinin [14] comparing scroll bars, dragging, a pop-up overview, and a pop-up overview containing a field-of-view indicator, showed that interfaces with a pop-up overview were significantly faster than both the scroll-bar interface and the drag interface.

Hornbæk, Bederson, and Plaisant [13] compared zooming interfaces with and without an overview found that, in contrast to previous results, subjects were faster without the overview. User preference was towards having an overview, though.

Combs and Bederson [4] compared four browsers for photographs: One thumbnail browser featuring a horizontal scrollbar, a zoom-and-pan browser, and two 3D browsers. Participants performed better with both 2D browsers than with the 3D browsers. Furthermore, they also made fewer incorrect selections with 2D browsers, and rated them higher. This contradicts previous results showing that scroll bars are slower compared to zoom.

In a study on the efficiency of fisheye browsers, Schaffer et al. [18] compared fisheye and zoom-and-replace interfaces on node-link diagrams representing a telephone network. For this task, the fisheye view was significantly faster than the zoom-and-replace.

Hornbæk and Frøkjær [12] compared linear, fisheye, and overview-plus-detail interfaces in order to find out how browsers affect the readability of electronic documents. They found that participants performed faster with the fisheye interface, but received higher grades with the overview+detail interface.

Contradicting the results indicating that fisheye would be faster than zoom, Donskoy and Kaptelinin [5] found that, in a comparison of scroll-bar, zoom, and fisheye interfaces, the lowest task completion times were achieved with the zooming interface. The fisheye turned out to be the slowest technique.

Gutwin and Fedak [9] made a comparison on small screens, finding that a fisheye interface was useful for web navigation tasks, and a two-level zoom for a monitoring task, while panning slowed down users, regardless of task type.
experiments, and a number of pilot studies, raised a lot of questions, which is the ground for the analysis presented in this paper.

A variety of image browsers was presented by Spence [19], and a description of image browsers is given in the image browser taxonomy by Plaisant et al. [17]. They describe image browsers in detail, and how different types of browsing tasks make use of different browser characteristics. The most extensive survey of research done within this field so far was made by Cockburn et al. [3]. In their summary, the authors state that it is difficult to provide guidelines from previous research, since there are many factors influencing the results. This paper provides an analysis of visual factors in image browsers. The aim is to give one explanation to why results from previous research differ so much, and highlight a few things to consider when conducting tests on image browsers.

III. MEASURING USABILITY

User tests involving image browsers are often conducted with task time completion, error rate, and ratings from subjects, and as measures of usability. In this paper we focus mainly on task completion time.

The time it takes to find a target image among others can be divided into three parts:

- The time it takes to recognize a possible target. (Seek)
- The time it takes to access the target. (Access)
- The time it takes to interpret the information once it has been accessed. (Interpret)

The time it takes to recognize a possible target depends on how easy it is to distinguish the target from the rest of the set [2], [21]. A click-to-expand technique is superior to most other techniques if the real target can be located in this phase. However, in a set of data where the target cannot be identified either by its characteristics, or by location, there might not be much to gain by clicking on each object, compared to using a zoom-and-pan or a fisheye function. This assumption has not yet been verified, though.

The time it takes to access the target, once it has been selected by the user depends on the technique used. Most techniques are divided into a sequence of operations: Move mouse and press button, or press button and move mouse, or move mouse, press button, and then move mouse again, etc.

The time it takes to interpret the information, once the image has been accessed, depends on type of data and, of course, on the type of task.

IV. LAYOUT AND PRESENTATION OF IMAGES

Image browsing is often discussed in terms of interaction technique or user task. Here we discuss the visual details, i.e. presentation and layout of images. There are a number of variables that could affect task completion time.

Visual factors can be divided into two groups. The first group consists of variables that are not limited by any other factor mentioned below:

- Display size
- Resolution
- Type of data
- Organization of images

The second group consists of layout and presentation parameters that are limited by the variables of the first group, and also may limit other variables within the group, and create confounding variables. For image browsers, these are:

- Number of image
- Size of images (i.e. size of representations)
- Spacing
- Appearance of images
- Display area usage (visual angle)
- Magnification ratio

In the following subsections, we discuss these variables in more detail. A dependency graph (Figure 8) illustrates the relationships between variables, and how each variable affects task completion time.

A. Display Size

The size of the display limits the largest possible size of images that can be viewed on the screen, and how large spacing between images can be. In combination with resolution, display size also affects the number of images that can be viewed at the same time and the magnification ratio between overview and detail view.

If the image cannot be viewed fully at a readable size, the time it takes to interpret the contents of an image could benefit from a larger display.

B. Resolution

Resolution affects the maximum number of images that can be displayed in the same area at once, and appearance of images. Whether the effect on the appearance of images is significant or not, depends on how important resolution is to recognition of various symbols. When it comes to text and numbers, readability is more affected by size of characters than by resolution [20], not necessarily meaning that users find size more important than high resolution.

C. Type of Data

Type of data is related to appearance of images. A simple symbol of a distinct colour and shape is clearly different from a complex image with many different colours and patterns. The time it takes to interpret information might be highly dependent on this factor.

In user tests aiming at comparing image browsing techniques, time spent in the interpretation phase should be as
short as possible. If the purpose is to find a certain target image, the content of the image should not be too complex. A simple and suitable task for this would be to identify a photo or some simple figure. If the task includes browsing a large image, a more complex image could be used in the test.

Complexity of an image varies with the amount of information represented in the image, and how it is presented. A detailed photo can contain a lot of detail that yet can be accessed all at once. A bus table on the other hands needs to be systematically searched in order to find certain information. When conducting a user test it is important to keep in mind that for complex tasks, performance can be affected by the level of user skills.

D. Organization of images

Organization of images is a matter of layout and sorting. Images can be laid out in various patterns (rows and columns, etc.) on the display. Sorting can be done by some characteristics, such as content or appearance of images (colour, shape, etc.). This factor could affect the visual search phase if the user recognizes the pattern.

E. Number of images

Number of images can, when the display is full, be increased by decreasing the size of the images. However, there is an upper limit for number of pictures that, when exceeded, will make the view useless without a proper browsing technique.

The largest number of objects that can be viewed at the same time is limited by size of images, spacing between images, and resolution. If resolution is high, more objects (but of smaller size) can be displayed.

F. Size of images

Size of images is, according to Fitts’ law [7], important to the time it takes to access an image. Besides affecting precision in the target access phase, size could also affect the time it takes to visually find a target or interpret the information.

Maximum size of images is limited by the number of images, and by spacing between images. If magnification ratio plays an important role (for instance when zoom interfaces are tested), the size of smaller representations of images will be determined by this factor. Or, magnification ratio will be determined by size of images, just as spacing and number of images. Size of images also affects appearance of images, since larger images are more easily recognized than smaller ones.

G. Spacing

Spacing affects, and is affected by, similar factors as size of images. Number of images, size of images, and the size of the display area limits maximum spacing. Even when using a fixed number and size of images, the display area usage, and thereby also the visual angle, will change if the spacing between the images is changed.

With Fitts’s law [7] in mind, it would be reasonable to assume that larger spacing would increase task completion time. However, the effects of spacing in an iconic interface were investigated by Everett and Byrne [6], and they found no evidence for this. It remains to be investigated how spacing affects other browsing techniques, and whether different techniques are efficient for different levels of spacing. Furthermore, there might be an interaction with other factors, such as display size (the effects of spacing could be larger on a larger display), type of data and appearance of icons.

In a comparative study, where users could use either a drag function or a click function in the same bifocal interface, it was observed that users seemed to prefer to click [11]. A possible explanation for this could be that objects separated by spacing encourage clicking.

H. Appearance of images

Appearance of images has an impact on the amount of scaling (magnification ratio) needed in order to interpret the content. Small images, icons, and other symbols are more easily recognized if they have different characteristics, such as colour, shape, and contrast. An image with a colour that is distinctly different from the rest of the set can be found just as quickly in a large set as in a small set [21], [22].

If all images have approximately the same appearance, and there is no text accompanying the image, there is nothing but location to go by in order to find a target image. And if the location is unknown, the user needs to go through the images one by one in order to find a target.

I. Display area usage

Display area usage refers to how much of the display is used, counting from the edges of the topmost and leftmost
images, to the lowest and rightmost images. This measure determines the visual angle. When either spacing, size of image, or number of images is changed while the other variables are kept constant, the visual angle changes automatically. (See Figure 4, 5, and 6.) This means that testing the variables above separately under equal conditions is difficult.

On smaller displays, the difference in visual angle between various levels of spacing is rather small. On a 50” display, however, the area to search can differ by several decimeters in each direction. This could affect both the search phase and the target access phase.

J. Magnification ratio

The magnification ratio, or zoom factor, is the ratio between the miniature image and the actual image. This factor is particularly important to consider when comparing a zoom interface to other browsing techniques. In a zoom interface, a large zoom factor means a longer distance to move along the z-axis, before a readable view is achieved. Furthermore, zooming in on a large high-resolution image could mean that the user loses too much context for the technique to be useful.

Panning and dragging operations are affected by the magnification ratio too. If the ratio is high, the actual area to pan in a zoomed in mode will be large, and the larger this area is, the more surface will lie "outside" the focus area. (Figure 7)

\[ t = a + b \log_2(1+D/W) \]  
(1)

The time it takes to access a selected image with a zoom interface, regardless of starting point, could be calculated by:

\[ t = a + b \log_2(1+D/W) + t_{zoom} + t_{pan} \]  
(2)

Either of \( t_{zoom} \) or \( t_{pan} \) could be 0, but not both. Hence, accessing one single image can not be done as quickly with a zoom interface as with an iconic interface.

We have tried to put focus on the fact that task completion time for image browsing could be affected by many factors. These factors could be affected by other factors, which makes it harder to say exactly what improves usability and what does not. Further testing is needed, but testing image browsers is not an easy task.

Assume we take into account a variable, factor A, in a usability test. Factor A is affected by another factor B, which we have not considered. Now, how do we tell what actually was the cause of the outcome of the test? If not all variables can be controlled, then it is harder to predict and explain the outcome, and results may be in conflict with results from other similar studies.

Each factor that has not yet been tested should be tested separately. Then interaction between various factors and browsing techniques should be tested. With this approach, it might be possible to clarify whether certain browsing techniques benefit from certain layout and presentation strategies.

However, there is more than layout and presentation that affects usability for image browsing. Plaisant et al. [17] described a number of different image browsing tasks and what browsing mechanisms were needed for various operations. It is reasonable to assume, just as Cockburn et al. [3] pointed out, that there exists an interaction between type of task and browsing technique. A logical way to continue would be to create a definition of properties of different tasks. If properties of different tasks are defined, it will be possible to vary those task properties and to investigate the possibility of an interaction between type of task and browsing technique.

Another category of factors to analyse is the hardware factors, such as type of input device, acceleration function, usage of buttons, etc.

Considering that user experience and task completion time sometimes differ, it is important to investigate subjective satisfaction further. Measures on mental workload can be added to the part of the analysis concerning subjective satisfaction, which would increase knowledge concerning the subject’s attitudes toward different browsing techniques.

Research that involves human beings is difficult. There is always a risk for confounding variables and unpredicted
behaviour. Thorough design and analysis of experiments will generate reliable results and a solid base to build further research on. Hopefully, this contribution is a step in that direction.

REFERENCES