Applied Ergonomics 45 (2014) 601–607

Contents lists available at ScienceDirect

Applied Ergonomics

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

Facilitating the comparison of multiple visual items on screen: The example of electronic architectural plan correction

Sylvain Fleury*, Éric Jamet

CRPCC, University of Rennes 2, Place du recteur Henri Le Moal, CS 24 307, 35043 Rennes Cedex, Rennes, France

ARTICLE INFO

Article history:
Received 30 November 2012
Accepted 8 August 2013

Keywords:
Plan correction
Spatial integration
Attentional guidance

ABSTRACT

This paper describes two experiments designed to (1) ascertain whether the way in which architectural plans are displayed on a computer screen influences the quality of their correction by humans, and (2) identify the visual exploration strategies adopted in this type of task. Results of the first "spot the difference" experiment showed that superimposing the plans yielded better error correction performances than displaying them side by side. Furthermore, a sequential display mode, where the second plan only gradually appeared on the screen, improved error search effectiveness. In the second experiment, eye movement recordings revealed that superimposition increased plan comparison efficiency by making it easier to establish coreference between the two sources of information. The improvement in effectiveness in the sequential condition was shown to be linked to the attentional guidance afforded by this display mode, which helped users to make a more thorough exploration of the plans.

1. Introduction

Large volumes of technical documents are currently interpreted automatically not just for archiving purposes, but also so that they can be modified using computer-aided design (CAD) software. These documents include land registry maps (Boatto et al., 1992), architectural plans (Lu et al., 2005; Ahmed et al., 2011), electrical diagrams (Ouyang and Davis, 2009), road maps (Chiang and Knoblock, 2011) and even music scores (Bainbridge and Bell, 2001). The two experiments described in the present paper were part of a collaborative research project to develop retroconversion software capable of interpreting scanned plans and reconstructing them in a format compatible with all the main types of CAD software.[1] Although this research will be generalizable to all types of technical documents, we chose to restrict our experiments to architectural plans. The software under development currently has a 9% retroconversion error rate for simple plans of this nature (Ghorbel et al., 2013). It will therefore be important for users to be able to identify mistakes. In some design activities, such as the creation of architectural plans, professionals often start off by producing a hand-drawn paper sketch (Bilda and Gero, 2006; Manolopoulou, 2005), which is then reconstructed in software, either manually or by means of automatic recognition. Whether we are talking about the manual copying of paper diagrams into CAD software, document retroconversion, or systems intended to beautify hand-drawn sketches of industrial parts (Ku et al., 2006), product designs (Orbay and Kara, 2010) or architectural plans (Elsen et al., 2012), it is vital that users inspect the end product to ensure that the system has not made any errors of interpretation. He must compare the two plans: the scanned plan and the one interpreted by the system. This type of "spot the difference" task, involving the mental integration of two distinct visual sources, raises the question of which is the best display format for ensuring that as many errors as possible are spotted.

1.1. Processing two visual sources

The cognitive processing of architects' drawings has been explored from several angles, including their 3D visualization (Yagmur-Kilimci, 2011), the different stages in their production (Bilda et al., 2006; Manolopoulou, 2005), and variations in design strategies according to expertise (Maycock et al., 2009). To the best of our knowledge, however, there has never been any research on the cognitive processes involved in comparing such drawings, the difficulties that may be encountered along the way and the best ways of minimizing them. There have, of course, been a great many studies of the parallel processing of visual information sources in the field of multimedia document comprehension (for a review, see Ginns, 2006). These studies have shown that when a document

* Corresponding author. Tel.: +33 2 99 14 19 58.
E-mail address: sylvain.fleury@uhb.fr (S. Fleury).

[1] This software is under development as part of the "MobiSketch" French National Research Agency (ANR) project (ref. 09-CORD-015).

0003-6870/$ – see front matter
http://dx.doi.org/10.1016/j.apergo.2013.08.004

consists of an image and a text that refer to each other, presenting each chunk of text next to the corresponding area of the image, rather than simply displaying the two side by side, enhances learning performances (e.g., Erhel and Jamet, 2006, 2011; Johnson and Mayer, 2012; Mayer, 2009; Sweller and Chandler, 1994). It is generally assumed that the beneficial effect of this spatial contiguity stems mainly from a reduction in the number of times that the gaze has to travel back and forth between the two sources. This to-ing and fro-ing is widely assumed to hinder individuals by forcing them to hold information in memory (Ayres and Sweller, 2005; Cierniak et al., 2009; Chang et al., 2011). Some authors, however, have claimed that spatial contiguity actually works by making it easier to establish coreference between the visual sources (Erhel and Jamet, 2006, 2011; Holsanova et al., 2008). This claim was recently confirmed by Johnson and Mayer (2012), who used eyetracker measurements to demonstrate that spatial contiguity leads individuals to make more eye movements from the text to the corresponding area of the image (and vice versa) than a separated format does. The principle of spatial contiguity has been extended from learning of text and pictures documents to comparative visual search by Bauhoff et al. (2012).

The research conducted by Kroft and Wickens (2002) on map comparison shed additional light on this subject. In their study, student pilots were shown two maps. One of them indicated the characteristics of the terrain (relief, power lines, etc.), while the other provided information about air traffic and weather conditions. Participants had to answer multiple-choice questions. Some of these questions (e.g., “What is the altitude of plane X?”) only required them to process one of the maps, but others (e.g., “Will plane Y fly over a lake?”) forced them to integrate information across both of them (Kroft and Wickens, 2002). Performances on the latter were better when the maps were superimposed, rather than being displayed side by side. Here once more, the beneficial effect of spatial integration was explained in terms of cognitive cost. Answering a question that concerned both maps involved three processing stages: information searching, reading and mental integration. When the maps were superimposed, the mental integration stage no longer required any cognitive resources (Kroft and Wickens, 2002).

1.2. Role of attentional guidance in document exploration

Correcting plans involves not just the comparison of multiple visual sources, but also their visual exploration. However, the nature of this exploration is determined by the image’s characteristics and does not necessarily take the form of systematic scanning. When individuals view an image displayed on a screen, for instance, they tend to fixate the centre of that image (Bindemann, 2010) and, more generally, any salient areas (for a review, see Grant and Spivey, 2003; Schütz et al., 2011). There are several methods for guiding the visual exploration of a screen, such as using motion to trigger visual pursuit (Dorr et al., 2010). This type of guidance was used in a study by Nickles et al. (2003), who found that a cursor moving across the screen improved visual search proficiency. According to the authors, it was the guidance afforded by the cursor that led to this improvement, by encouraging a more exhaustive exploration of the image. Tracking down errors in plans relies on precisely this sort of exhaustive exploration. Given that retroconversion software interprets documents one element at a time, one solution would be to use the realtime display of this gradual process to guide attention. This on-screen retroconversion would constitute a form of sequential display. A great deal of research has shown that when a fresh item appears on the screen, it immediately captures the viewer’s attention, triggering a saccade to that item and the allocation of processing resources (Abrams and Christ, 2003; Craig et al., 2002; Godijn and Theeuwes, 2002; Hillstrom and Chai, 2006; Ludwig et al., 2008). Furthermore, several studies have reported improved comprehension of a multimedia document when it is displayed sequentially (Bétrancourt et al., 2001; Jamet et al., 2008; Mayer and Chandler, 2001). One explanation for these results is that sequential presentation avoids perceptual and cognitive overload by ensuring that there is never too much information on the screen at any one time. Another is that it ensures that the document is explored in a more coherent order (Bétrancourt et al., 2003; Jamet et al., 2008).

2. Experiment 1

We began by conducting a “spot the difference” pretest, where participants had to compare two plans displayed side by side. The results of the pretest highlighted the inherent difficulty of this task, in that only 33% of participants managed to detect every single error. Research on multimedia documents, albeit ones combining textual and pictorial information, suggests that spatially integrating the two plans would have improved performances (e.g. Ayres and Sweller, 2005; Erhel and Jamet, 2011; Mayer, 2009). Furthermore, as mentioned earlier, Kroft and Wickens (2002) concluded in their study that integration makes it easier to answer questions requiring the combined processing of two maps. These parallels with maps and multimedia documents led us to predict that plan integration would reduce the need to hold information in memory, thereby boosting task effectiveness.[2] Our second prediction was that the attentional guidance effects observed for sequentially presented multimedia documents would also be observed when the second plan in our comparison task only gradually appeared on the screen. We postulated that the sequential display mode would facilitate plan comparison by minimizing the visual search process and proposing an order of exploration which, if the participants adhered to it, ensure that each and every feature of the plan was checked. By making the mistakes easier to spot, it would ultimately improve the effectiveness of the retroconversion software prototype.

2.1. Material and method

2.1.1. Participants

This first experiment was administered to 54 participants (19 men and 35 women), students and young professionals, recruited from the basis of voluntary testers of the Observation Laboratory on the uses of information and communication technology (LOUSTIC). The youngest was aged 18 years and 9 months, and the oldest 31 years and 3 months. The participants’ mean age was 23 years and 2 months (*SD*: 38 months).

2.1.2. Material

Participants performed the spot the difference task on an Asus Eee Slate with a 12.1″ screen, examining three different plans interpreted by our retroconversion software prototype. This prototype is capable of recognizing all the different symbols used in hand-drawn plans, but for the purposes of our experiment, it only had to recognize walls, which it showed as black lines, doors (red boxes) and windows (blue boxes). The source document appeared on the screen at the beginning of each test, but participants were not allowed to start circling the errors until the retroconversion process was finished (30–60 s, depending on the plan’s

[2] In accordance with ISO usability standard 9241, *efficiency* referred to the time it took to complete the task and *effectiveness* to its successful completion (participants’ ability to locate all the errors).