Supporting shape reinterpretation with eye tracking

Conference or Workshop Item

C. M. Herr, N. Gu, S. Roudavski, M. A. Schnabel, *Circuit Bending, Breaking and Mending: Proceedings of the 16th International Conference on Computer-Aided Architectural Design Research in Asia,* 101–110.

SUPPORTING SHAPE REINTERPRETATION WITH EYE TRACKING

IESTYN JOWERS,[1] MIQUEL PRATS,[2] NIEVES PEDREIRA,[3] ALISON MCKAY[4] and STEVE GARNER[5]
1. University of Leeds, Leeds, UK, i.jowers@leeds.ac.uk
2. The Open University, Milton Keynes, UK, m.prats@open.ac.uk
3. University of A Coruña, A Coruña, Spain, nieves@udc.es
4. University of Leeds, Leeds, UK, a.mckay@leeds.ac.uk
5. The Open University, Milton Keynes, UK, s.w.garner@open.ac.uk

Abstract. It has been argued that reinterpretation is an essential process in design generation and idea exploration. However, computational design tools, such as computer-aided design systems, offer poor support for shape reinterpretation, and as such are not well suited to ideation in conceptual design. One of the key difficulties in implementing computational systems that support shape reinterpretation is the issue of interface – how can a user intuitively guide a system with respect to their interpretation of a designed shape? In this paper, a software prototype is presented that uses an eye tracking interface to support reinterpretation of shapes according to recognised subshapes. The prototype is based on eye tracking studies, and uses gaze data and user input to restructure designed shapes so that they afford manipulation according to users' interpretations.

Keywords. Eye tracking; shape interpretation; computer-aided design; design generation; design exploration.

1. Introduction

This paper presents a software prototype that uses an eye tracking interface to support the creation and manipulation of designed shapes in the generative and explorative stages of conceptual design.

Conceptual design, in a number of creative professions, involves the creation, exploration, and development of design shape alternatives, and is typically supported using freehand sketches. The ambiguity of sketches supports the cognitive process of image reinterpretation which is central to effective shape exploration and development (Schön and Wiggins, 1992). Commercially available computational design tools, such as computer-aided design systems, are typically not well suited to conceptual design activities because the support they offer for the reinterpretation of shapes is poor; they are instead typically used during detail design and design definition.

When discussing reinterpretation in design it is important to distinguish between interpretation of the *semantics* of a shape, and interpretation of the *syntax* of a shape (Goldschmidt, 1991). The first is concerned with assignment of meaning and involves cognitive processes that are difficult, if not impossible, to formalise; while the second is concerned with the structure of the geometric elements that define design descriptions. This research is concerned with supporting syntactic interpretation, which is also a vital element of any potential support for semantic interpretation, as discussed in Section 5. The software prototype presented here makes use of emerging eye tracking technologies to support dynamic reinterpretation of designed shapes according to recognised structures, and is based on experimental studies in which the gaze of participants was recorded as they looked at and manipulated shapes. The resulting system is an advancement towards what Saund and Moran (1994, p. 175) view as an ideal system for design exploration, and is intended to

> read the user's mind (his visual system in particular) and synchronize with whatever image entities the user happens to perceive as significant.

Eye tracking technology has the potential to substantially reduce the cognitive overhead needed for designers to interface with computational design systems. The prototype presented here is a first imagining of the possibilities, and is intended to allow designers to focus on their design activity, with the software as a tool that supports without disturbing the creative flow.

2. Background

The ability to interpret and reinterpret shape is a fundamental skill in visually creative activities such as conceptual design. It has been observed that, when sketching, designers often produce series of ideas that are, in places, inevitably ambiguous and open to reinterpretation. These design alternatives are explored visually in a search for patterns and associations that lead to new avenues of exploration. Schön and Wiggins (1992) describe this as a 'seeing-moving-seeing' process where seeing a sketch can result in its reinterpretation according to emergent forms or structures, and this in turn informs the development of future sketches. Exploration in this way typically involves the recognition and transformation of shapes in sketches, such as overall outline shapes or the embedded parts of shapes, so-called ‘subshapes’ (Prats et al, 2009).

Despite the importance of reinterpretation in design exploration, it is not readily afforded by the current generation of computational design tools. For example Goel (1995) and Stones and Cassidy (2010), present comparative studies in which designers undertake conceptual design tasks using either sketching or commercial computational design tools. In both of the studies it was found that participants readily use reinterpretation in their design exploration if sketching, but not when using computational tools. However, Stones and Cassidy observe that reinterpretation did take place cognitively when computational tools were used, although this was not evidenced in the creation of new solutions. They suggest the reason for this is because, when participants were using computational tools, they were looking for accuracy in their design concepts and until a form closely resembled their mental picture they were unable to progress to alternative interpretations. Lawson and Loke (1997) propose a more pragmatic reason and suggest that development of computational design tools has placed too much emphasis on graphical representation techniques. As such the resulting tools are unable to support processes essential to explorative design, including the process of shape reinterpretation.

A key difficulty in supporting shape reinterpretation is the problem of developing an interface that allows the user to specify their current interpretation of a designed shape so that the system can afford the required manipulations. Conventional selection techniques such as pointing, clicking, or encircling shapes with a mouse are not always practical because of ambiguity that arises due to overlapping interpretations. Research into this problem has resulted in a range of approaches that enable the manipulation of designed shapes according to recognised structures. For example, Saund and Moran (1994) present a WYPIWYG (What You Perceive is What You Get) drawing system that enables perceptual interpretations of a digitally sketched shape to be specified and manipulated according to simple pen-based gestures; Gross (2001) presents the ‘Back of an Envelope’ system, a drawing program that uses standard pattern recognition techniques to automate the recognition of emergent subshapes in a digital sketch; Jowers et al (2010) present an approach based on a shape grammar formalism (Stiny, 2006), in which shape replacement rules are applied to identify and manipulate recognised subshapes in a design.

A major limitation to each of these approaches lies in the cognitive overhead needed to interface with the systems. For example, in Saund and

Moran's system, users are required to learn gestures that specify particular interpretations of a shape; in Gross' system users are required to trace over shapes of interest in order to specify their interpretation; in Jowers et al's system users need to define shape replacement rules before shapes can be manipulated. In each of these, additional effort is needed to interact with the system to specify a particular interpretation of a shape and, no matter how small this effort is, it can result in an interruption to the creative flow of the user. A more intuitive, dynamic system, one that better supports a cognitive process of 'seeing-moving-seeing', would offer real benefits in avoiding the need for users to explicitly define their interpretation of designed shapes. To this end, eye tracking technology presents itself as a potential interface for computational design tools.

Previous research has explored the application of eye tracking as an alternative drawing interface, to replace traditional mouse and keyboard input, e.g. Hornof et al (2004). Here, eye tracking is used as an additional interface, augmenting traditional input. In the software prototype, gaze data is used in combination with mouse-input to reveal a user's interpretation of a designed shape at a particular moment in time, and the software responds by affording manipulation of the recognised subshapes.

3. How do users look at shapes?

In order to inform development of the software prototype, a series of eye tracking studies was conducted with the aim to build an understanding of how gaze data can be used to support shape reinterpretation in a computational design tool.

Gaze data, consisting of scan paths and points of visual fixation, can reveal much about how shapes are viewed and interpreted. The studies reported here attempted to identify gaze patterns that could be used to distinguish between shapes that are attended to and those that are simply looked at. Attention to a shape is a necessary process that precedes interpretation (Fu et al, 2006), and is identified by focus on one particular shape in an image, whilst ignoring others that are also present. This act of attending is linked to the act of selective looking but not necessarily to the act of recognising.

In the studies, participants completed tasks in which they were encouraged to employ different visual strategies, including free-viewing, shape search, and attention to specified shapes. For example, in the search tasks, participants were asked to search images composed of shape primitives for specified shapes, and to vocalise when they had found the target shape. An example of the images presented to the participants is given in Figure 1, along with an example of the gaze data collected. Here, the participant was asked to find the arrow highlighted in Figure 1a, a task made more difficult due to the necessity to undertake a figure-ground reversal.

During the studies, gaze data was collected using a Tobii X120 – a non-intrusive eye tracker that allows for some head movement. Verbal data was also collected facilitating a comparison between gaze patterns that relate to seeing shapes and gaze patterns that relate to attending to shapes. Analysis of the data identified patterns of viewing consistent with attending to shapes. It was found that, when participants were seeing shapes without attending to them, the scan paths followed no pattern of note but tended to be loosely concentrated around the centre of the image, with the majority of visual fixations taking place inside individual shape primitives, as illustrated in Figure 1b. However, a dominant pattern emerged when shapes were attended to, with a majority of scan paths and fixations concentrated around the centre of gravity of the attended shape, as illustrated in Figure 1c. This confirms the findings of Vishwanath and Kowler (2003), who report that participants attending to simple shapes naturally fixate on the centres of gravity.

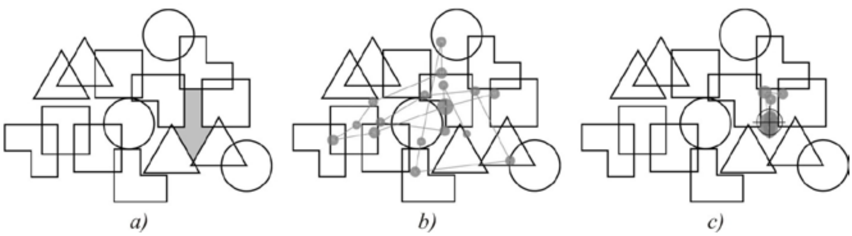

Figure 1 – Stimuli and corresponding data from a shape search task

The studies also explored how gaze data reflects interpretation of ambiguous shapes. This involved tasks in which participants were asked to fixate on specified subshapes of an ambiguous shape. For example, in one task participants were shown the five-point star in Figure 2, and were asked to attend to specified subshapes, examples of which are illustrated in Figure 2a. The aim of these tasks was to determine whether or not it is possible to infer a viewer's interpretation of an ambiguous shape by comparing their visual fixations with the positions of the centres of gravity of potential subshapes of the shape. As with the shape primitives, it was found that when participants were seeing a shape without attending to any particular subshape the gaze data had no pattern of note. But, a dominant pattern emerged when subshapes were attended to, with a majority of scan paths and fixations concentrated around the centre of gravity of the attended subshape, as illustrated in Figures 2b and c.

However, it was also found that if a shape is composed of subshapes that have centres of gravity that are close together (within a distance of approximately 50 pixels), then it is not possible to unambiguously determine which of the subshapes is being attended to based purely on gaze data. For example, the gaze data illustrated in Figures 2b and c could not conclusively identify which of the two subshapes the participant was attending to because their centres of gravity are too close. Instead, it is only possible to identify subshapes that were *potentially* attended to.

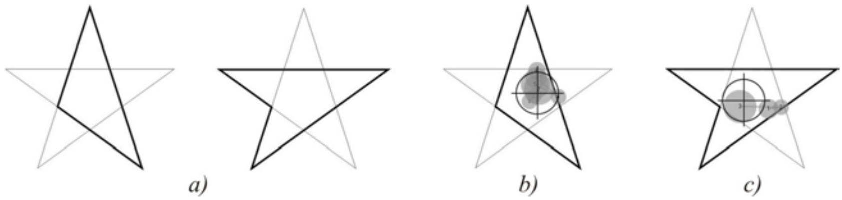

Figure 2 – Stimuli and corresponding data from a shape interpretation task

4. Supporting shape reinterpretation

The results of the eye tracking studies were used to inform development of a software prototype, implemented using the Tobii Software Development Kit. Key to implementing the prototype was the identification of potential interpretations of a designed shape. Also, there was a need to understand users' interpretation of designed shapes based on potentially ambiguous gaze data and mouse-based input. In this section, these issues are discussed, and the methods used to address them are presented. The resulting software prototype is illustrated with an example showing how the user is able to manipulate subshapes of interest in a designed shape.

4.1 IDENTIFING POTENTIAL SUBSHAPES

Central to the use of gaze data for determining a user's interpretation of an ambiguous shape is information about the centres of gravity of potential subshapes embedded in the shape. The brute force approach to determining this information is to calculate all possible subshapes and then calculate their centres of gravity. For example, the subshapes can be calculated combinatorially by considering sets of lines that are embedded in the shape. However, this approach is computationally very expensive, and will result in consideration of subshapes that are very unlikely to be considered by a designer manipulating the shape, e.g. subshapes composed of single lines.

The approach used here is more selective, and is similar to that implemented by Gross (2001). In the software prototype, users are able to build up a library of shapes that they find interesting. The library is populated in two ways:

firstly, any shape that the user explicitly draws is added to the library; secondly, the user can trace over any interesting emergent subshapes and so add them to the library. Shapes in the library are compared to the current designed shape using a method of subshape detection, as described by Trescak et al (2009), and every similar subshape that is detected is made available for selection by the user.

4.2 RESOLVING AMBIGUITY IN SUBSHAPE SELECTION

As discussed in Section 2, when manipulating recognised subshapes in a designed shape, mouse-based methods of selection are not always practical, especially if different interpretations of the shape overlap. This is because the geometric elements that are used to construct shapes can be shared by different interpretations, as illustrated in Figure 3a, where mouse-based selection cannot unambiguously determine whether the user is selecting the triangle or the square. Input from an eye tracker can be used to determine which subshape the user is attending to at a particular moment but, as discussed in the previous section, if subshapes have centres of gravity that are close to each other then it is not possible to unambiguously determine which specific subshape is being attended to. This is illustrated in Figure 3b, where gaze data cannot unambiguously determine if the user is attending to the outer or the inner square, because they share a centre of gravity.

A combination of the two methods of input can serve to cancel out much of the ambiguity that can arise. In combination, gaze data can resolve the ambiguity in mouse-based input, and mouse-based input can resolve the ambiguity in data from an eye tracker. This is illustrated in Figure 3c where the combination of the two methods means that the user can unambiguously select and manipulate the subshape of interest, the inner square.

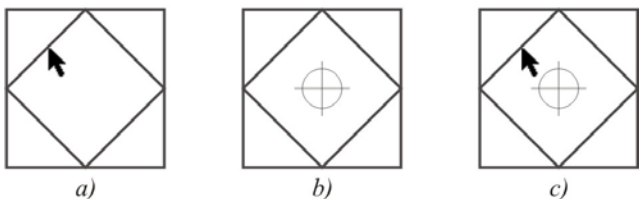

Figure 3 – Ambiguity in shape interpretation

4.3 SUPPORTING MANIPULATION OF RECOGNISED SUBSHAPES

These methods of identifying potential subshapes and resolving ambiguity in subshape selection have been implemented in a software prototype, the user-interface of which is illustrated in Figure 4. The user-interface consists of a large drawing area (on the left hand side), a shape library (on the right hand side), and various eye tracking tools, such as the dialog in the bottom right corner that reports the current status of the gaze data being recorded. In the software, shapes are defined as sets of line segments, and subshapes are defined according to subsets of these. Users can manipulate constructed shapes by translating the end points of individual line segments, or by selecting and translating defined subshapes.

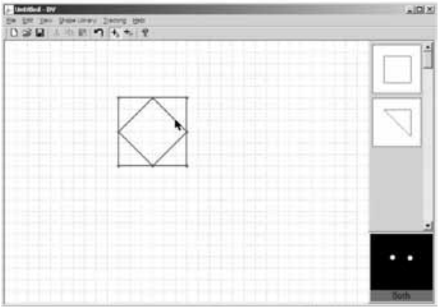

Figure 4 – The software prototype

Initially, subshapes are defined according to how a shape is constructed but as shape exploration proceeds, and the shape library is populated, the constructed shape can be reinterpreted according to recognised subshapes, which can then be selected and manipulated. For example, in Figure 4 the shape was constructed as two squares, and a square and a triangle have been added to the shape library. Accordingly, in Figure 5, this designed shape is manipulated by translating recognised square and triangle subshapes.

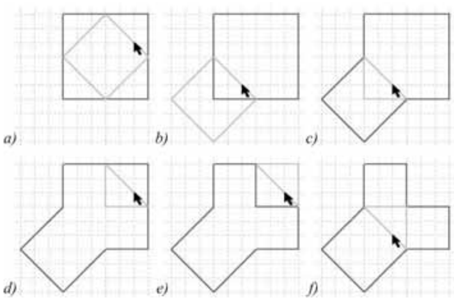

Figure 5 – Manipulating subshapes of interest using the software prototype

Note that the mouse-based input remains constant throughout the example, with the user always selecting the same line segment. However, because the user is attending to different subshapes according to different interpretations of the designed shape, it is possible to manipulate different parts of the shape. This is clearly illustrated by comparing Figure 5d and 5e, where the user attends to different triangles and consequently is able to select and manipulate those triangles, in turn.

At each stage of the explorative process the centres of gravity of the detected subshapes are used in combination with mouse-input and gaze data to determine the current interpretation of the designed shape – if a user is attending to a subshape that is similar to a shape in the library, and if that subshape is selected using mouse-input, then it is made available for manipulation through the application of an identity shape rule (Stiny 1996). Such rules are of the form $A \rightarrow A$, where A is a recognised subshape, and application to a designed shape S, results in the shape $(S - A) + A$. Identity rule application serves to restructure the designed shape S so that the recognised subshape A is defined and is available for manipulation.

5. Discussion

This research is an exploration of the two-way conversation between the designer and the design representation – the maker and the made. Schön (1988) argues that design does not conform to an objectivist philosophy, where things are what they are independently of how they are interpreted. Instead, he argues that design is constructionist. Designers construct their knowledge of a design problem as they explore potential solutions. Central to this process are *gestalts*, coherent wholes that are defined by designers' interpretations of the geometric elements that compose design representations. Gestalts enable designers to reason about design problems, and are not fixed. The same set of geometric elements can be reconstructed as many different coherent wholes, and a designer often shifts gestalt during design exploration. The research presented here has explored how a computational design system can support such gestalt shifts, and it has proposed a method for computationally supporting the constructionist philosophy of design.

Gestalts result from semantic interpretations of shapes, involving cognitive processes that are beyond the scope of this research. But, manipulation of a shape according to a particular gestalt is only possible if the structure of the shape, i.e. its syntax, is compatible. Therefore, in order for a computational design tool to support gestalt shifts it is essential that syntactic interpretation of shapes is afforded. Here, syntactic interpretation is supported using an eye tracking interface. The combination of gaze data with traditional user input serves to dispel much of the ambiguity that can arise with respect to the user's intended manipulation of a designed shape. As a result, the software prototype is able to support reinterpretation of designed shapes according to alternative sets of embedded geometric elements that reflect the user's current perception of the shape.

It is a difficult task to develop a computational system that can recognise different interpretations of shapes, and the solution presented here uses a compromise whereby the users themselves specify potential interpretations based on recognised subshapes, but in a manner that is as non-intrusive as possible. It is anticipated that this will allow designers to focus on their design activity, with the design software as a tool that supports without disturbing the creative flow.

Acknowledgements

This research reported in this paper was carried out as part of the Designing with Vision project (http://design.open.ac.uk/DV), funded by The Leverhulme Trust.

References

Fu, H., Chi, Z., and Feng, D.: 2006, Attention-driven image interpretation with application to image retrieval, *Pattern Recognition*, **39**(9) 1604-1621.

Goel, V.: 1995, *Sketches of thought*, MIT Press.

Goldschmidt, G.: 1991, The dialectics of sketching, *Creativity Research Journal,* **4**, 123–143.

Gross, M. D.: 2001, Emergence in a recognition based drawing interface, *Proceedings of Visual and Spatial Reasoning in Design II,* Sydney, Australia, 51–65.

Hornof, A. J., Cavender, A., and Hoselton, R.: 2004, EyeDraw: A system for drawing pictures with eye movements, *Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility*, Atlanta, GA, 86–93.

Jowers, I., Hogg, D., McKay, A., Chau, H. H., and de Pennington, A.: 2010, Shape Detection with Vision: Implementing Shape Grammars in Conceptual Design, *Research in Engineering Design*, **21**(4), 235–247.

Lawson, B., Loke, S. M.: 1997, Computers words and pictures, *Design Studies,* **18**, 171–183.

Prats, M., Lim, S., Jowers, I., Garner, S., and Chase, S.: 2009, Transforming shape in design: Observations from studies of sketching, *Design Studies* **30**(5), 503–520.

Saund, E., Moran, T.: 1994, A perceptually supported sketch editor, *Proceedings of ACM Symposium on User Interface Software and Technology*, Marina del Rey, CA, 175–184.

Schön, D. A., Wiggins, G.: 1992, Kinds of seeing and their functions in designing, *Design Studies,* **13**(2), 135–156.

Schön, D. A.: 1988, Designing: Rules, types and worlds, *Design Studies,* **9**(3), 181–190.

Stiny, G.: 1996, Useless rules, *Environment and Planning B,* **23**(2), 235–237.

Stiny, G.: 2006, *Shape: Talking about Seeing and Doing*, MIT Press.

Stones, C. and Cassidy, T.: 2010, Seeing and discovering: how do student designers reinterpret sketches and digital marks during graphical design ideation?, *Design Studies*, **31**(5) 439–460.

Trescak, T., Rodriguez, I., Esteva, M.: 2009, General Shape Grammar Interpreter for Intelligent Designs Generations, *Proceedings of Sixth International Conference on Computer Graphics, Imaging and Visualization*, Tianjin, China, 235–240.

Vishwanath, D., and Kowler, E.: 2003, Localization of shapes: eye movements and perception compared. *Vision Research*, **43**(15) 1637–1653.