Providing blind people with access to technical diagrams

Paul Blenkhorn[1], David Crombie[2], Sijo Dijkstra[2], Gareth Evans[1], Bláithín Gallagher[3], Cornelius Hagen[4], Mirko Horstmann[5], George Ioannidis[5], Alasdair King[1], Mark Magennis[3], Anne-Marie O'Neill[6], Helen Petrie[6], Christoph Schlieder[4], John Wood[3]

[1]Department of Computation, UMIST, Manchester, UK, [2]FNB, Amsterdam, Netherlands, [3]National Council for the Blind, Dublin, Ireland, [4]University of Bamberg, Germany, [5]TZI, University of Bremen, Germany, [6]Centre for Human Computer Interaction Design, City University, London, UK

Keywords :Blind, diagram, tactile, interface, access

Abstract: This paper describes the operation and evaluation of a series of prototype systems that have been developed (as part of the EU-funded TEDUB project) to give blind people access to diagrams in well-defined technical domains: software engineering diagrams (UML), architectural plans and electronic diagrams. The prototype systems focus on how information (including logical structure and spatial relationships) can be conveyed to a blind user through the use of structured hierarchies and a variety of input and output modalities (including sound (2D and 3D) and force-feedback joysticks). The paper describes the context of the TeDUB project, the architecture of the prototype tools and the final system, and gives detailed evaluation results based on trials with over 20 users. The evaluation results show that connectivity and relative position can be effectively conveyed, but that the relationship between logical and spatial structure needs further development.

Introduction

Some professionals use diagrams specific to their knowledge domains. Diagrams are used for a number of reasons, including: the facilitation of rapid communication of ideas and information through the ability of readers to examine and interpret visual information with great speed; the storage of technical reference information in a format that has an agreed common meaning; and, finally, the storage of information in working documents, utilising them as external memory aids. Such diagrams generally use a well defined vocabulary of symbols and meanings agreed as a standard within the profession, for example Universal Modelling Language (UML) used in the software industry (UML 2003). The use of consistent vocabularies means that semantic information can be expressed by simple diagram components: for example, the existence of an arrow between two items in a UML diagram indicates a defined relationship between the two items.

We classify diagrams as being one of two types. In the first type spatial layout is essential to the understanding of the diagram (e.g. maps). In the second type spatial layout may aid the understanding of the diagram, but it is not essential, and diagrams of this type can be drawn in many different orientations and spatial layouts. This type of diagram is generally of the form of a connected network and includes many forms of electronic circuit diagram and most of the UML diagram types. In this type of diagram the relative positioning of the content of a diagram can impart information even if it is not absolutely necessary: for example, the grouping of a number of electronic circuit components may indicate to a sighted user a common electronic function performed by the components as a unit. Bennett and Edwards (1998) did not find blind users utilising this compositional approach, but their study used tools that communicated this information to the blind user in any other way except through text. Conversely, the compositional structures defined in software engineering Data Flow and N^2 diagrams have been found to be of use to blind users for comprehension when a dedicated user interface was provided to deliver this information in the JUSTIN and KEVIN systems (Blenkhorn and Evans, 1998; Petrie *et al*, 1996)

The TeDUB system is designed to allow blind users to read technical diagrams. Its distinguishing features are: it assumes no special preparation of the diagram; that it uses cheap and common non-specialist technology; and that it allows users to use diagrams as external memory aids, organising and annotating them so they are working documents rather than simply fixed sources of information.

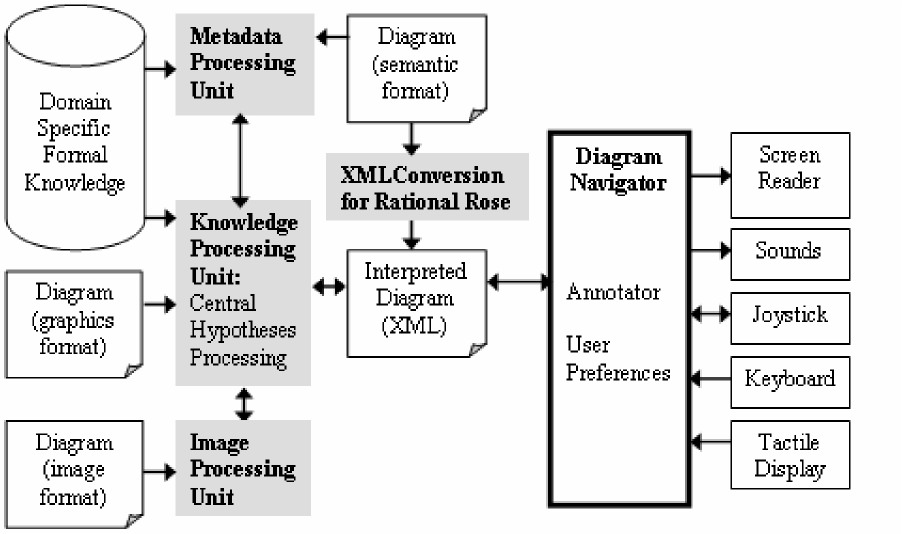

Figure 1: The architecture of the TeDUB system.

System architecture

Diagrams come in different formats with different levels of abstraction. At the lowest level, a diagram can be stored as a *raster graphics* picture, a collection of pixels. Typically, a scanned picture is represented this way. The format at the next level is *vector graphics,* which is typically produced by graphics programs. Diagrams are stored as sets of geometric shapes like lines, arcs or circles. A third type of format consists in storing the semantic content of the diagram which is independent of its geometric representation. Metadata like annotations belong to this level. CASE-Tools like Rational Rose store diagrams in a format of this kind, usually augmented with geometric information. The TeDUB-system is designed to handle diagrams regardless of which of the three levels they represent.

As Figure 1 shows, the Diagram Interpreter (modules shown in grey) builds an interpreted representation of the diagram. Here, the main challenge is how to relate visual information to semantic information. For example, if a user is not familiar with digital circuits, a diagram of a digital circuit will seem to them to be a meaningless collection of lines. This shows that extracting the diagram’s semantics from geometric information requires knowledge of the specific domain. The relation between visual information and semantics is ambiguous: the same geometric information can represent different semantics, e.g. a tree of boxes can stand for an organigram or a family tree. The converse is also true: the same semantics can be realized in different geometric terms, for example electoral votes in a pie- or bar-chart.

The knowledge processing unit

Diagram Interpreter consists of several modules. Its core is the knowledge processing unit which is based on an inference engine. It operates on a network of hypotheses which represent parts of the diagram, ranging from low abstraction geometric objects to high abstraction metadata. In typical operation, upward inferences generate from given geometric hypotheses further hypotheses on the semantic level, and process them incrementally until a semantic description of the whole diagram is found. Additionally, top-down conclusions make it possible to integrate information coming from metadata and, to a certain extent, to reconstruct badly recognized parts of a diagram. A central blackboard-data structure serves as interface between the units, especially the metadata-processing- and the image processing-unit.

The image recognition unit

Bitmap images, like those acquired through a scanner, consist of two-dimensional matrices of pixels. In order to extract vector level information from bitmap-images, the image recognition module passes through a number of processing steps. After a pre-processing procedure that deals with various possible distortions in the image, groups of pixels that belong to one and the same object are identified, a process known as segmentation. Whereas humans can easily separate, for example, pixels belonging to a connection line from pixels belonging to a component in an electronic diagram by using their domain knowledge, an image recognition system is only able to classify groups of pixels by their geometric properties. In the subsequent vectorisation step, these groups are processed and described by their position, shape and other attributes. As a result, the image recognition module produces primitive hypotheses like “straight line”, “curve” or “rectangle” that are then passed on to the knowledge management unit.

The user interface and diagram access

The user interface, Diagram Navigator, works only with XML files in the Interpreted Diagram TeDUB format, produced by Diagram Interpreter, or by an XML transformation conversion from diagrams output in XML from the Rational Rose UML tool. It can perform some rudimentary editing functions: bookmarks and annotations can be applied to the diagram and saved, and viewed in a hierarchical structure applied. It is not, however, a diagram creation application: it is intended to allow diagram use, not production.

Diagram information can be accessed in three different ways. Firstly, all the content of the diagram, which does not relate to spatial location, is available through a hierarchically-structured presentation. The choice of a hierarchical organisation is based on the following: the ability to hide detailed information behind contextual summary information helps blind users quickly to orient themselves when first encountering the diagram, as suggested in Bennett & Edwards (1998); hierarchical document navigation is familiar to many users through webpages and other structured documents such as DAISY audio books; and finally it allows compositional structure to be represented to blind users by the aggregation of child items into parent items that correspond to visual groups obvious to sighted users. Secondly, the user can explore the diagram spatially to discover the connectivity and spatial location of the items in the diagram. This spatial layout is orthogonal to the hierarchical layout: spatial connections between items will not respect hierarchical level. Thirdly, the text content of the diagram can be accessed and searched directly. The differentiating factor between this approach and the two exploratory approaches is the directed, focused search for information: the user may be interested in specific information content, not comprehension of the diagram (for example, the answer to “Does this electronic circuit contain a capacitor?”) or the user may already be familiar with the diagram from previous exploration and simply want to obtain information that the user knows is there already (for example, “What is the value of that capacitor I found while exploring?”).

User interfaces and modalities in Diagram Navigator

The simplest user interface supported is the use of a keyboard for input and simple text output in standard Windows controls. This maximises accessibility through screen readers, allowing users to use their familiar and preferred screen reader to output to voice synthesis or Braille display as required. This simple interface allows for navigation of the hierarchical structure by use of the cursor keys in the style of Windows Explorer, which evaluation of prototype systems indicated to be familiar to many users. It also borrows from web browser design with a "breadcrumb" function (a "back" button) and a "home" at the top of the hierarchy. These are intended to help the user to orient him or herself within the hierarchy and to navigate the structure with confidence. Direct access to the text content of the diagram is supported by a search function, user annotation of items in the hierarchy, persistent bookmarks, and limited editing of the hierarchical structure, all accessed through the keyboard and text output.

A number of different audio interfaces can be used individually or jointly as the user prefers. The first supports basic usage of the user interface. Simple action sounds play to provide information on user actions, such as "gone up a level". Examples of these can be found in many applications, for example the DHANI system developed by the EU ACCESS project (Petrie *et al*, 1997; Morley *et al*, 1998): they are simple and provide immediate feedback on the success or failure of user actions. These are intended to allow the user to build a mental model of how the program works and how to control it and obtain information from it.

The second way to use sound is with context sounds, similar to those developed by Brewster (1998). These provide absolute positioning information on the user's location within the structured diagram hierarchy. They are intended to play continually and can combine pitch, tune, timbre, tempo and instrument as discriminating factors.

Finally, the system uses 3D sound, generated by a standard PC soundcard. The context sounds can be played relative to the position of the items they identify in the diagram. The location of items around the user's current item can be indicated by a progressing "radar sweep" around the user indicating the position of connected items. Otherwise the 3D capabilities are used in conjunction with the joystick to access the spatial content of the diagram and will be discussed with the joystick functions.

Joysticks have been used in other systems for communicating spatial information to blind people, including graphs (Yu, 2001) and maps (Schneider & Strothotte, 2000). These systems have tended to use specialised haptic devices such as the PHANToM force-feedback joystick (SensAble 2003). The TeDUB system uses commercially-available force-feedback games joysticks, which are cheaper but more limited in function (Johansson, 1999). Two modes of operation are supplied: in the first, the joystick is used to indicate the direction of connections to the user by force-feedback, and, in the second, the user can indicate a direction and be informed of any connected item. The user can also move between items and therefore explore the whole diagram through spatial navigation with the joystick. 3D sound reinforces the joystick function by spatialising connected items around the user when the joystick indicates their presence: the plane of joystick action matches the plane of sound spatialisation around the user.

User trials

Trials were performed on over twenty blind and partially-sighted users with the system, ranking the components on a 1 to 5 scale. The hierarchical model navigated with the keyboard was well received (mean score 4.75), possibly due to its similarity to the familiar Windows Explorer model of navigation. Other simple text-based navigation and query functions were scored highly: the search function scored 4.04, although this was limited to searching for known items in the diagram by name; the home function scored 3.96. Again, both these functions are to be expected to be familiar to users from other applications.

The action sounds were well received, which contrasts strongly with the poor response to the context sounds (2.04). This may be due to the poor design of the context sounds: it is difficult to construct sounds that are informative without being obtrusive and irritating. Conveying spatial information with sound alone was not well-received: playing an absolutely-located sound to indicate the current item's location scored 2.75, and playing the sounds of surrounding items to indicate spatial positions relative to the current item scored 2.94. Users gave similar scores to the sounds' abilities to communicate distance: 2.60 and 2.25 respectively. Isolating the sound functions from the reinforcing joystick functions and the active process of exploration and navigation was not successful.

Using the joystick and supporting sounds to convey spatial information was better received. Users preferred moving the joystick to query for neighbouring items (3.94 and 6 users) to having the joystick use force-feedback to communicate neighbours directly to the user (3.44 and 3 users). They also found it easy to explore the spatial layout of the diagram by moving from item to item using the joystick (3.94 and 3.96 for the two approaches). This process may support the user in building a mental model of the connectivity and positional information of items in the diagram. Access to the non-spatial functions through the joystick buttons was appreciated, providing control without having to remember keyboard command key combinations (4.17). However, the users expected some relationship between the hierarchical structure and the spatial structure, and the lack of any such relationship confused them. This suggests that designing for some correlation would be a positive improvement, but the tree structure cannot be made to resemble the spatial layout in any meaningful fashion in a general case.

Conclusion

The TeDUB system has implemented a user interface to communicate spatial and content information from a diagram to a blind user that has performed well in user trials. Work will progress over the remainder of the project on integrating this interface with image analysis and file import to provide access to technical diagrams for blind people.

References

Bennet, D. J. and Edwards, A. D. N. (1998) "Exploration of non-seen diagrams", in Bennet, D.J. and Edwards, A. D. N. (eds) *Proceedings of ICAD '98 (International Conference on Auditory Display)*, British Computer Society, 1998.

Blenkhorn, P. and Evans, G. (1998) "Using speech and touch to enable blind people to access schematic diagrams", in *Journal of Network and Computer Applications* vol 21, 1998, pp 17 - 29.

Brewster, S. (1998) "Using Non-speech Sounds to Provide Navigation Cues" *ACM Transactions on Computer-Human Interaction*, 5(2) 1998, pp 26 - 29.

Johansson, A. J. and Linde, J. (1999), "Using simple force feedback mechanisms as haptic visualization tools", *Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference (IMTC99)*, Volume 2, Venice, Italy, May 1999, pp 820 – 844.

Morley, S., Petrie, H., O'Neill, A.-M., and McNally, P. (1999). "Auditory navigation in hyperspace: design and evaluation of a non-visual hypermedia system for blind users". *Behaviour and Information Technology*, 18(1), pp 18 - 26.

Petrie, H. Obee, J., Blenkhorn, P., and Evans, G. (1996) "User-based evaluation of a system for access to computer aided software engineering (CASE) tools for blind software engineers", *Proceedings of the 5th International Conference on Computers Helping People with Special Needs*, Linz, Austria, July 1996, pp 375 – 380.

Schneider, J., and Strothotte, T. (2000). "Constructive Exploration of Spatial Information by Blind Users", *Proceedings of ASSETS 2000*, ACM Press, November 2000, pp 188 – 192.

SensAble (2003) Website http://www.sensable.com/products/phantom_ghost/phantom.asp (10 April 2003)

UML (2003) Website http://www.uml.org/ (10 April 2003)

Yu, W., Ramloll, R., and Brewster, S. (2001). "Haptic Graphs for Blind Computer Users" in *Proceedings of the First International Workshop on Haptic Human-Computer Interaction*, Springer-Verlag, Berlin, 2001, pp 41 - 51.