CASE STUDY OF VOICE CONTROL OF AUTOCAD

Corine Bickley*, Christopher S. Jones*, Janet Wang,
Hong Z. Tan*, Robert Ebert*, David M. Horowitz*

Massachusetts Institute of Technology
*also at Department of Rehabilitation Medicine, Tufts University School of Medicine

ABSTRACT

A system for voice, head pointer, and switch control of AutoCAD for a quadriplegic architectural designer has been developed. The user was trained on AutoCAD and each of the assistive devices in term, using a "Wizard of Oz" paradigm. Command sets within the voice recognizer were utilized in order to increase recognition accuracy.

INTRODUCTION

Voice input for controlling software which can be used in vocational tasks has been shown to be feasible by a number of researchers (see, for example, (2), (3) and (4)). The focus of our current work has been to develop a number of specific voice-input applications for individuals with disabilities. These applications include the creation of reports by voice with the Kurzweil Voice Report system and of engineering drawings by voice and head pointing. The goal in each case has been to develop systems and to provide training that will result in employment of the individuals in jobs which are consistent with their backgrounds, interests, and vocational skills.

In this report, we describe one case study. The user of the system is an architectural designer who needs to produce both reports and engineering drawings. Our goal has been to develop a system which integrates on one workstation the capabilities of voice reporting and hands-free drawing.

BACKGROUND

Reports, such as customer quotes and bills of material, are commonly used in the day-to-day operation of a small business, such as our client's architectural design company. The Kurzweil Voice Report (KVR) system was programmed to create several voice-activated templates for the standard reports which our client has used routinely in his business. The voice reporting part of this project is described in detail in (3).

Engineering drawings, specifically architectural house plans, are also needed by our client. Architectural plans are one kind of engineering drawings which are standardly generated using computer-assisted design (CAD) programs, such as AutoCAD.

CAD software allows a user to create drawings which are precise and can be modified easily. CAD technology supports the design of generic modules, which can be manipulated in various ways to create a large number of house plans from a small set of drawings of individual components, such as walls, supports, closets, etc. AutoCAD primitives, such as lines and arcs, are sufficient for creating 2D floor plans and elevations, and also details such as windows and siding. An alternative to drawing the plans using only primitive geometric objects is to use one or several of the readily available macro packages which can supplement the drawings which an individual designer creates. Window and door CAD packages, for instance, are available directly from suppliers of these parts. Other detailed drawings, such as kitchen cabinets and bathroom fixtures, are available in packages such as DCA and CADKIT.

CAD programs are usually controlled by keyboard and mouse commands, but recently have been shown to be accessible by voice and head pointing. Burnett and his colleagues have developed systems for voice-only control of AutoCAD and phone handling (2). The success of this approach has allowed training programs such as those at the Seattle Resource Center for the Handicapped to familiarize users with these sorts of systems.

SUBJECT

A quadriplegic architectural designer (JEM) has been a client in our program of demonstrating the use of voice in vocational rehabilitation. JEM has retained good control of his voice and head movement, despite his C-5 quad injury in 1988 which left him with minimal arm movement and little fine motor control. He is not able to use a keyboard for prolonged periods of time, and is thus unable to control software such as AutoCAD by keyboard commands or mouse movement.

JEM has an excellent, in-depth knowledge of the field of architectural design. He operated his own business for 13 years prior to his accident, and completed over a thousand plans for individuals and small businesses. JEM had recognized the usefulness of CAD for architectural design in 1988, and in fact had ordered a license for AutoCAD just prior to his accident.

SYSTEM DESIGN

The first step was to select system components which would provide JEM with the functionality he needed to prepare house plans and which could be integrated with AutoCAD. We chose the combination of the FreeWheel head pointer (from Pointer Systems), the BUG voice recognizer (from Command Corp.), and a customized button switch (as a substitute for a mouse click). The FreeWheel is a passive target system which uses a small clip-on reflector which can be easily attached to the subject's glasses and a small sensor device which can be positioned on top of the computer monitor. There are no wires which must be connected to the computer; thus the user is freer to move about the office without disconnecting and to position him- or herself in front of the computer.

The BUG voice recognizer was selected because of its robust recognition and relatively large vocabulary (up to approximately 200 words in up to 32 command sets). The use of command sets allows the recognizer to compare a specific voice token to only a subset of the entire lexicon. The user is allowed to structure the command sets in order to maximize recognition accuracy. For instance, the words "N" and "end" are acoustically similar, but can be separated by being included in different command sets. Thus, when the user needs to label a drawing, and thus might need to say the letter "N", the command set named Spell can be turned on with the voice command "Spell" and the active vocabulary will include all of the letters of the alphabet, but not the word "end". In order to reduce the confusability of the letters such as "B", "D", "T", etc., the standard communication words such as "bravo" and "tango" were used for many of the letters of the alphabet.

A large button was chosen as the clicking device for JEM because he is able to use gross arm movements to control devices such as joysticks and toggle switches. A single voice command would have been appropriate for a subject for whom pushing a button is not convenient (1).

TRAINING

The next step was to train JEM in the use of AutoCAD, the FreeWheel head pointer, and the BUG voice recognizer. We chose to separate the tasks, training JEM on the use of each device one at a time, in order to simplify the cognitive load during the training sessions. We used a "Wizard of Oz" training protocol, in which the subject's speech and pointing were recognized by a human and translated into computer input. After JEM had mastered the use of AutoCad, the FreeWheel pointer and the large button were incorporated in training sessions in place of the human. Then the computer-based voice recognizer, the BUG system, replaced the human listener. A standard office setting was used for speech training in order to simulate a realistic environment. After voice token training was completed, JEM began to use the complete system to create house plans. During training sessions, he worked with a drawing which he had previously drawn by hand prior to his accident. In using such a familiar drawing, we were able once more to minimize the cognitive load which training imposed.

RESULTS

Using the BUG recognizer, the FreeWheel pointer, and the button switch, JEM successfully recreated various details of the house plan. JEM acquired skill with the FreeWheel pointer and the button switch quickly. We found, though, that the FreeWheel pointer needed frequent recalibration and was sensitive to JEM's sitting position.

Recalibration was easily accomplished by moving the pointer to the extremes of the screen (left to right, and top to bottom). Head pointing proved to be somewhat tiring; frequent rests and stretches of the neck muscles were used to minimize the effects of long periods of use of the pointer.

Voice token training on the BUG system was straightforward. Recognition accuracy seems to be sensitive to specification of command sets, as was expected. JEM is currently working with five command sets (draw, modify, spell, and tools, as well as a set of commands which is never off).

DISCUSSION

Qualitatively, we (the subject and the research team) are encouraged by the preliminary results. We are all confident that JEM will be able to restart his business and to use voice control to create the house plans needed for his customers. The FreeWheel head pointer and button combination for point specification worked well functionally. Head pointing appears to be a tiring activity after several hours, and it will probably require some time for a user to develop proficiency. The BUG voice recognizer served our task well; the use of several command sets appears promising. We anticipate fine tuning the use of command sets over the next several months. The recognizer is somewhat sensitive to microphone placement, as expected. JEM easily became proficient at adjusting the microphone for good recognition. Room noise during voice training and recognition proved to be a minor complication. Because the BUG system can be easily retrained on specific words, and because it behaves gracefully when confronted with a word which it does not recognize (e.g., it continues to wait for a word which can be recognized), JEM was able to retrain words as needed. We anticipate that JEM will find that point specification by voice (i.e. specifying coordinates using words for the numbers and dimensions) is more useful than by head pointer because architectural drawings require exact locations to be specified. Thus, we expect that any difficulty which JEM would experience after many hours of head pointing simply will not occur.

FUTURE WORK

We plan further work in the area of evaluation of JEM's use of the combined system, both quantitative and qualitative. We hope to focus on issues of quantifying JEM's speed in creating drawings by voice compared to an able-bodied user creating drawings by keyboard and mouse, and also on identifying factors which affect the usability and user satisfaction of voice CAD in actual vocational situations.

ACKNOWLEDGEMENTS

This project was supported by the U.S. Department of Education Rehabilitation Services Administration Grant No. H128A01011-92.

REFERENCES

1. Burnett, J.K. (in preparation) A Networked, Voice and Head Movement CAD Workstation for Quadriplegics.

2. Burnett, J.K., Klabunde, C.R., and Britell, C.W. (1991) Voice and Head Pointer Operated Electronics Computer Assisted Design Workstations for Individuals with Severe Upper Extremity Impairments. RESNA '91, Proceedings of the 14th Annual Conference. RESNAPRESS, Washington, D.C., pp. 48-49.

3. Horowitz, D.M. and Tan, H.Z. (1992) Utilizing Speech Recognition Technology to Increase Productivity - Case Studies. RESNA '92, Proceedings of the RESNA International '92 Conference. RESNAPRESS, Washington, D.C., pp. 61-63.

4. Tan, H.Z. and Horowitz, D.M. (in preparation) The Kurzweil Speech Recognition System - Is It For You?