Department of
Veterans Affairs

Journal of Rehabilitation Research and
Development Vol. 31 No. 4, November 1994
Pages 345–354

# Voice output reader for displays on video cassette recorders and other domestic products

**D. Gareth Evans, BSc and Paul Blenkhorn, BSc**
*Department of Computation, University of Manchester Institute of Science and Technology (UMIST),
Manchester, M60 1QD, UK*

**Abstract**—The increasing use of electronic displays in domestic products can pose great problems for blind people wishing to use microwave ovens, video cassette recorders, hi-fi systems, etc. One method of trying to solve these problems is to provide a handheld "display reader" that can translate the alphanumeric and symbolic information presented on the display into speech. A prototype system has been developed and evaluated. User requirements for this device and the problems of the image sensor, the image processing method, and the user interaction with the system are also discussed.

**Key words:** *assistive technology, blind, computers, consumer electronics, image processing, speech synthesis, visual impairment.*

## INTRODUCTION

The two main areas identified as posing difficulties for blind people (1), are mobility (getting from A to B) and reading. Although the latter was originally perceived as being concerned with access to printed material, this now includes access to information generally, and, in particular, information stored and presented in systems employing computer technology. Considerable efforts have been made to enable blind people to access computer systems, and a significant area of recent interest has been providing access to computer systems employing a graphical user interface (GUI). These have resulted in a number of solutions predominantly for IBM-compatible equipment (2,3); it should be noted, however, that the first commercially available GUI access system was for the Macintosh [outSpoken from Berkeley Systems (4)].

In parallel with developments in desktop computing, there also has been an increasing tendency to use electronic displays and keypads in products such as video cassette recorders (VCRs), washing machines, and microwave ovens. As these displays become more commonplace, they can pose a serious obstacle to blind people operating these devices, which traditionally could be modified by adding tactile markings to the dials and switches of the products. This difficulty was highlighted recently by the World Blind Union Research Committee, who listed, among their priorities for technical research and development for visually disabled persons, a liquid crystal display (LCD) reader for numeric displays with speech output for less than 100 dollars (5). The work reported here has been undertaken to address this need by producing a prototype system that is suitable for further development to produce a small, handheld device with a self-contained power supply and capable of speaking the information found on typical consumer products with electronic displays.

### Accessing Domestic Products

Toward the end of 1992 a survey of major United Kingdom retailers was undertaken to determine the state of the market in domestic products with electronic displays. The products chosen for the survey ranged from VCRs and microwave ovens to digital clocks and calculators. Displays on personal computers and organizers were not considered. Of particular concern were "leading-edge" products; these were examined with the objective of improving our understanding of both current and future trends

---

Address all correspondence and requests for reprints to: D. Gareth Evans, C38, Main Building, Technology for Disabled People Unit (TDPU), Systems Engineering Group, Department of Computation, University of Manchester Institute of Science and Technology (UMIST), P.O. Box 88, Manchester, M60 1QD, UK.

in display technology and techniques. Displays were classified in terms of the underlying display technology, how the technology was used to present different data items on the display, and the structuring of the information on the displays.

### *Types of Display Technology*

Three different types of display technology were found: 1) liquid crystal displays (LCDs); 2) light-emitting diodes (LEDs); and 3) gas plasma displays (GPDs).

LCDs are low-contrast, dark-on-light displays and are gradually becoming less common on domestic products in general, due to their limited viewing angles and relatively low contrast. However, they are common on battery powered devices, due to their low power consumption.

LEDs are bright and clear, light-on-dark displays, generally found on many types of domestic equipment. They have generally been superseded by GPDs, but are still found generally on simple, low-cost displays, such as alarm clocks.

GPDs were found to be the most common form of display used. Like LEDs they are bright and clear, but can be constructed to provide higher resolution, and therefore to produce more detailed displays.

### *Presentation of Information*

The basic display technology can be used to present information in a number of ways. These include: 1) 7-segment displays; 2) 16-segment displays; 3) display symbols; and, 4) bit-mapped displays.

1. Seven-segment displays have the display segments arranged in such a way as to form numerals and a small subset of the alphabet.
2. Sixteen-segment displays are similar to seven segment displays but have an extra nine segments, which permit a better representation and a wider range of alphanumeric characters. These were found to be the most common displays in use.
3. In display symbols technology, a part of the display is constructed to present a particular symbol that can be "on" or "off." For example, the dots used to indicate a.m. or p.m. on some digital clocks, or the symbols that indicate the state of a VCR (i.e., play, pause, etc.). These symbols, referred to as "static elements" are found on almost all displays.
4. In bit-mapped displays a wide variety of symbols may be represented from a matrix of pixels. At the moment these are only present on top-of-the-range products, but it is anticipated that they will increase in usage over the next few years. It is important to note that this type of display has significantly different characteristics from the other display types in that the symbols can move in the display area.

### *Structure of Information*

**Figure 1** shows a display that is typical of those found on VCRs. This figure illustrates some of the terms used to describe and structure information found on displays. The display is considered to be made up of a number of *fields*, which are in turn made up of *elements*, which in turn are made up from *cells*. Cells are the fundamental units on the display that can be independently turned on or off. In the example, cells would include the recording icon, the "SUN" area, or any of the individual segments of the digits. Cells are grouped together to form elements. Elements may be alphanumeric or static symbols, which are made up of one or more cells. In the example above, elements made up of single cells include the recording icon and the "SUN" area; elements made up of multiple cells include the individual seven segment digits. Fields are sets of elements that, as a whole, form the fundamental units of information to be spoken. In the example above, four distinct fields can be identified as follows:

- a time field, consisting of four numeric elements and a separator (":")
- a tape counter field, consisting of five numeric elements
- a day-of-the-week field, consisting of seven static elements ("SUN," "MON," . . . "SAT"), which under most circumstances, would have a single element lit.

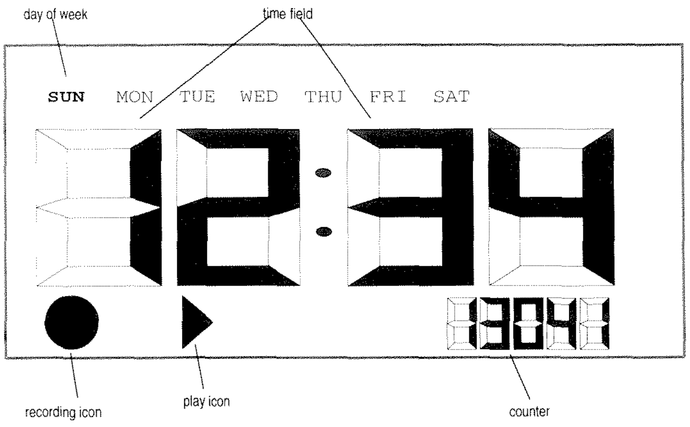

**Figure 1.**
An example of a display on a typical VCR.

- a state of system mechanism field, (i.e., a symbol representing RECORD, PLAY, FAST FORWARD, REWIND, PAUSE, etc.). In this case, between zero and three elements would be displayed. For example: zero in the case of the tape not being played; one for the tape being rewound; two in the case of "RECORD" and "PLAY"; and three in the case of a paused recording.

It should be noted that at any one time a user is typically interacting with one of the fields. In addition, on some systems information is scattered around the product (e.g., the "recording on" light on some VCRs is well away from the other information), or when information is not displayed on the product (e.g., with "on screen" VCR information). The work discussed here only deals with systems that indicate the system state in a single display panel.

### Issues in Developing a Display Reader System

To address the difficulties of accessing domestic products, the authors propose a low-cost display reader. The system would be a small, handheld device with a self-contained power supply. The user would place the system in front of a display, possibly locating it in some form of jig. The system would then present information appropriate to the current display and system status to the user with synthetic speech. Users would have selective control over the information relayed to them by a set of user-interface buttons. The display reader system, attached to an electronic system with the relevant type of display, is shown in **Figure 2**.

In developing such a system the authors had to address two issues: the usability of the device and the technological issues pertaining to the "reading" of the display.

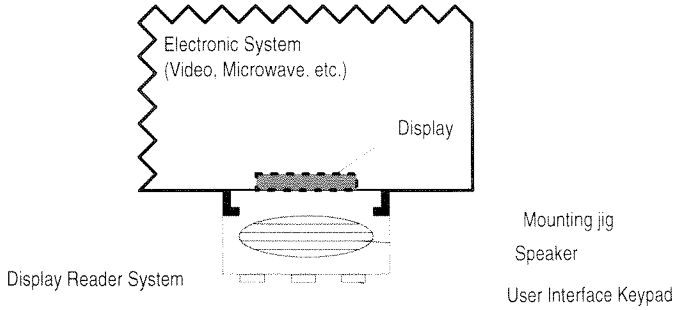

**Figure 2.**
Plan view of the display reader system attached to an electronic system.

The usability issue breaks down into three areas: device portability, the mechanical properties of the device, and the user interface.

A portable device must be of reasonably small size and also be cordless (i.e., it must be battery powered and hence the system design must minimize power consumption). To satisfy these requirements the system's chip count must be minimized, probably by using standard-cell application specific integrated circuit (ASIC) technology. If a sufficient volume of sales (i.e., in the order of hundreds) could be generated, this technology would significantly reduce the cost of the system.

The mechanical properties of the device impact greatly on its usability: users must be able to accurately position the device in front of the product's display and maintain this position for the period in which the product is in use. This prototype system locates the reader using a jig that significantly simplifies the alignment issues.

The user interface is a critical part of the design in such a system and may well determine whether the system is destined for "life in a cupboard" or is to become a valuable aid for the visually disabled person. The user interface is described in detail below.

The technological problems associated with the system are concerned chiefly with the derivation of robust image processing algorithms suitable for the application. The image processing algorithms are required to work with all commonly used display technologies and a wide range of display types. Other problems are concerned with the identification of a low-cost, low-power, small-sized image sensor having the appropriate level of spatial and contrast resolution for this application and the system design to support a good and useful user model.

From the above, one can see that the system needs to be highly integrated and to use some form of ASIC-based design. Yet there are two areas that require considerable levels of experimentation: in the image processing area, to prove display technology independence of both the image sensor and the display reader algorithms; and in investigating the usability of the system, both in terms of the user interface and in the packaging of the product. This need for experimentation is rather at odds with the high-cost development route associated with producing a highly integrated product. One could resolve the necessary experimental issues by using computer-based modeling techniques; where executable, prototype models of the system running on a computer system may be evaluated for their image processing capability and the suitability of the proposed user interface. Such a technique has been described (6), albeit directed at another class of applications. These techniques are effective methods of assessing design issues before developing highly integrated (so-called deeply embedded) systems. However, in this project the development of a physical prototype was chosen.

The primary reason for choosing a physical prototype is so that the usability of the system can be evaluated by undertaking controlled experiments with visually disabled users. The authors felt that a physical prototype would be more appropriate for this exercise. A further, more pragmatic reason concerns the fact that this work was undertaken by a group of master of engineering degree students working on an assessed project of the type where rules demand that skills in both software and hardware engineering be demonstrated.

The project has been carried out with full consideration of a final usable product. Although size, power, and cost constraints have not been directly addressed in the construction of the prototype, they have been fully accounted for by the system design. The design is technology independent and may be reused in an embedded system implementation.

## User Model

Although the fundamental requirement for this system is to enable a user to operate consumer electronic products, as has been stated above, it is critical that this is achieved in a way that results in the easy use of the product. The objective has been to make a reactive system in which the user operates a product in a manner that closely resembles that used by someone without a visual disability, and where the speech output from the system is given automatically when appropriate for a given context. This is illustrated below by considering user interaction with a typical VCR.

To assist in meeting the goals for the display reader, lessons can be learned from computer speech access systems for blind people. These "screen readers" enable blind people to operate standard applications running on IBM-compatible personal computers. There are many thousands of blind users of such systems around the world. Screen readers have developed from systems in which users could ask for information about areas of the computer screen (7), such as the current character, word, or line, into sophisticated systems in which configurations or profiles are constructed for given applications and that provide application-specific knowledge (8). Although users have the facilities to produce their own configurations, configurations for the major applications typically are made available by the companies who produce the screen readers. The importance of making such systems configurable and adaptable is detailed elsewhere (9), where the objective is stated as making the speech access system provide a seamless interface between application and user.

An important issue that should be noted when considering speech systems is that the access needs to be balanced with the quantity of information. "Minimum speech with maximum information" (10, p. 537) has been our philosophy in the design of the user interface for this system.

The approach for the display reader has been to construct a configuration for each display, which is set up so that, in most circumstances, the system automatically responds with an appropriate speech message as the user interacts with the consumer appliance. This is achieved by providing the facility to label fields as:

- *read field on change*, where the contents of the field, or one of a set of speech messages, will be spoken when a change occurs in that field
- *read field continuously*, where the field status will be repeatedly reported at a specified frequency
- *read once*, which is generated by a user request.

For each of these options the spoken text can be derived from the display field, a previously stored text string, or a combination of the two. For example, "The time is 10 p.m." where the information "10" is extracted from one field on the display, and the "p.m." is extracted from another field. In addition, an option is provided for a user to manually step through and examine different fields.

These capabilities can be illustrated by describing a user interaction with a typical VCR, where the user is to record a program that is about to start on the television set. Initially the user will attach the reader to the jig on the recorder and will insert a tape into the machine. The user now presses REWIND to ensure that the tape is at the start. The display changes and a "<<" symbol appears on the display that the display reader notices as a change and speaks the message "rewind." When the system stops, the display symbol "<<" is cleared and the system will announce "stop." At this point, the user may wish to reset the program counter and presses the COUNTER RESET button. This is not spoken and the user will need to manually operate the reader to step to the COUNTER field and read the contents. (The counter change is not automatically read on change as this would be very distracting in most circumstances). The user then presses RECORD and PLAY which again change the display and the system would respond by announcing "recording."

## System Architecture

### *The Logical View*

The system architecture is presented in **Figure 3** using structured analysis notation (11). This is a logical view of the architecture and is independent of implementation. The actual prototype system on which initial experimentation has been based is shown in **Figure 4**.

The way in which the system operates is briefly described with reference to **Figure 3**. The Image Sensor is placed in front of a display. In essence, the sensor is a video camera that continually passes video information to the system. The Capture Image and Image Store are essentially a hardware frame grabber, which digitizes the video signal and stores it as a pixel array in the Image Store memory. Although not implemented in the prototype system, where standard components have been used, there is a strong possibility that in the final system the Capture Image process may perform some elementary image processing operations (e.g., thresholding) "on the fly," as pixels arrive, and before they are placed in the Image Store. Images are captured when requested by the Capture Image flow; completion of image capture (i.e., an updated Image Store) is signaled by the Image Captured flow.

The image is then processed to determine the semantic information currently presented by the display. This is accomplished in two stages: Process Image, which operates on the raw pixel information to determine which areas of the display (cells) are illuminated; and Analyze Image, which operates on the list of cell states to determine the meaning of the information and whether this should be spoken. Both processes reference the Display Configuration Store. This store is divided into two parts: a static part, which contains configuration information for the particular display (e.g., the fields in the display, the locations of elements and of cells in the fields, the text associated with each field, etc.), and a dynamic part, which is a reflection of the current state of the user preferences and determines, for example, which set of fields should be examined, the format of the text to be spoken, etc. The image processing algorithm is described in detail later.

The System Control process determines when a frame should be captured and processed, and what, if any, speech should be passed to the user. A Capture Image message is sent either in response to some user command that explicitly requests that a display field or fields, be read, or is generated based upon the current operating mode of the system, as recorded in the Display Configuration Store. For instance, when the system is configured to "read field on change" (see the section above on User Interface), frames must be periodically grabbed at a rate exceeding one frame per second. The User Interface carries out text-to-speech synthesis on strings of text and processes commands received from the user via the keypad.

The image processing functional units, Process Image and Analyze Image, rely upon having access to the unique characteristics of the display held in the Display Configuration Store; this is the static part of the store mentioned above. This information is used by Process Image to locate

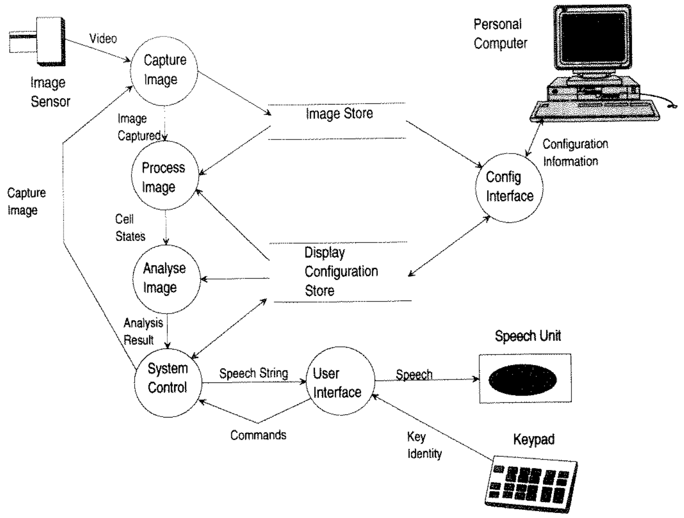

**Figure 3.**
The system architecture.

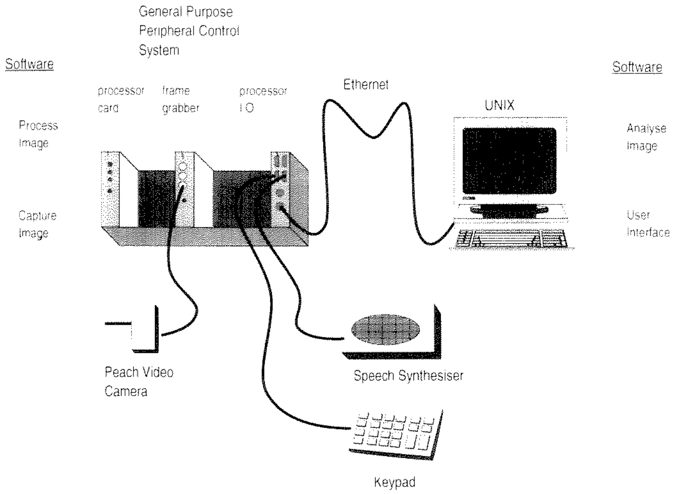

**Figure 4.**
The display reader prototype system.

fields within the Image Store and to determine the format of those fields. Analyze Image uses this information to determine the text strings to be spoken. Thus, for each display to be read there is a configuration entry in the Display Configuration Store to supply this information to these processes. This information is required each time the display reader reads a particular display and is therefore held in nonvolatile memory. There must be some way of capturing the configuration information for each display; which, it is anticipated, will have been produced by a sighted operator. This is supplied through the Config Interface, which is a process that needs access to both the Image Store and the nonvolatile part of the Display Configuration Store. The issues relating to system configuration and the role of the sighted operator are discussed later.

### The Prototype System

The prototype system, on which experimental work has been carried out, is briefly described in this section, and is illustrated in **Figure 4**. The display reader system prototype has been developed using a General Purpose Peripheral Control System (GPPCS) interfaced to a UNIX workstation. This environment (12) allows an application programmer to develop prototypes of embedded systems using the features of the UNIX operating system. In general, applications programs are written for, and reside in, the UNIX workstation. In this prototype system the Process Image, Analyze Image, and User Interface were developed in the UNIX system and interfaced to the relevant input/output (I/O) devices through peripheral dialogue library (PDL) functions. These functions present a high-level interface to the application software. For instance, the PDL function for the speech synthesizer will accept strings of text as parameters. The PDL functions manage communication with the GPPCS and call the relevant device driver library (DDL) in the GPPCS; at this level of abstraction the speech synthesizer is controlled by an RS232 driver procedure. The scheduling of tasks within the GPPCS is controlled by a small kernel that runs on a 68030-based processor card and controls the I/O devices over a VME bus.

In most circumstances, the communication overhead between the GPPCS and the UNIX workstation is insignificant. However, in the display reader prototype the Process Image process has relatively low CPU overhead but operates on the whole captured image. In this case, the overhead associated with shipping the entire image to the UNIX workstation and processing it there was about four times that of running the Process Image procedure within the GPPCS and transferring the cell states to the Analyze Image process running in the workstation.

The prototype system acts as a sound platform for the development of the software for the system, which could be ported with little modification into an embedded system product. The prototype does not reflect the hardware architecture of a portable system. However, the prototype system was used to evaluate the most critical hardware component, namely, the image sensor.

The image sensor must be of low cost and small size; have minimal power requirements and sufficient spatial and contrast resolution for the application; and provide acceptable quality images over a wide range of display types. The first three of these requirements are largely satisfied by the "Peach camera" manufactured by VLSI Vision (Edinburgh, UK). This is a 312 × 287 pixel image sensor array with on-chip circuitry to deliver a full-format composite video output signal. The ASIC Image sensor (13) may be purchased as a die, mounted on a printed circuit board (PCB) or packaged as a miniature camera, complete with wide angle lens. The latter option was used in this project; the entire video camera measures 3.3 × 3.5 × 2.7 cm. The camera is of relatively low cost (around £50, approximately, U.S. $75, per unit in volume) and has good power characteristics, dissipating less than 150 mW. The camera, therefore, appears well suited to this application. Evaluation of the camera was carried out in the prototype environment using a number of display types; the results of this evaluation are summarized in the results section below.

## METHOD

### Image Processing

A robust image processing algorithm is required to read the wide range of display types available in the domestic market. As the display reader is required to be held in a jig immediately in front of the display, ambient lighting conditions can be controlled. For emitter displays (LEDs and GPDs), the emitted light is uniform from all areas of the display and these displays have sufficient contrast between lit and nonlit areas for simple adaptive thresholding techniques to be used. Nonemitter displays (i.e., LCDs) are available in two types: backlit, where the emitted light from the display gives sufficient illumination for treatment in the same way as emitter displays; and nonbacklit, where a light source must be provided by the display reader unit to illuminate the display. In the latter case, careful design of the unit is needed to ensure reasonably uniform lighting of the complete display and the elimination of reflection from the surface of the display. In the prototype system jigs have not been used, the camera has been held in front of the display and ambient light has been used to illuminate nonemitter displays. Further work is necessary in jig design to develop satisfactory lighting conditions.

Standard optical character recognition (OCR) techniques are not applicable for reading product displays. In order to determine the meaning of symbols (static elements) within the display, the algorithm must refer to a set of configuration details for the display. For instance, a single dot on a display, which in another field presents time, may indicate that the displayed time is p.m. (or conversely a.m., there is no standard convention). An entry must exist in the Configuration Display Store for this information to be correctly interpreted. Thus, the characteristics of displays force the system to have a configuration entry for each display. However, once this principle is accepted, it actually makes the design and implementation of the image processing method rather more simple than if conventional OCR techniques were used. This is because the configuration entry can indicate to the Process Image function the set of cells, elements, and fields in the display, their configuration, and their precise position in the display. In the limit, this means that, if the camera is aligned precisely in the same way as when the configuration details were obtained, the image processing task becomes a reasonably simple task of checking various areas of the screen to see whether they are emitting light.

The prototype system relies on precise alignment (i.e., for the system to work the camera must be in the same position relative to the display as when the configuration details were captured). In the final system such alignment would be difficult to attain. Each domestic appliance in the user's house would need a jig to hold the display reader in front of the display. Locating the camera accurately in the jig each time is not a difficult problem to solve; however, it is likely that such a system may be difficult to use and may include potentially costly financial overheads both for manufacture and installation. However, the idea of a jig is appealing because it gives the user hands-free operation of the display reader; ensures that the display is entirely within its field of view; and holds the image sensor at a fixed distance from the display, thus maintaining a fixed character size. Therefore, it is assumed that a jig will provide some form of coarse, first-order alignment of the display within the field of view. For instance, the jig may be a number of Velcro™ strips around the product's display that interlock with similar strips on the display reader. A software process will align the current display image with that captured in the configuration store. Given that the precise nature of layout of the display is known, this is not a difficult task. In fact, a simple alignment program was developed for the prototype system; this worked well in principle but was defeated to some extent by the "fishbowling" introduced into the image by the camera (see the results section below).

The image processing method is now briefly described; it assumes that accurate alignment is maintained.

The configuration entry for the display identifies the fields on the display, the type of elements, which make up the (fields) and the precise location of the cells, which form the elements. For the low-level image processing algorithm the primary unit of concern is the element. These were defined earlier and can be one of four formats: display symbol, 7-segment character, 16-segment character, or dot matrix character. By knowing the size and format of the element, the exact position, shape, and size of the cells that form the elements can be determined. Thus for a 7-segment display, each of the seven cells that make up the display can be tested to determine if they are illuminated. Illumination of a cell is determined by thresholding the image, counting the number of illuminated pixels within the cell area, and comparing the value to some defined threshold. This process is illustrated in **Figure 5**.

The list of cell states for each element is returned, together with the field and element identifiers, for subsequent processing. From the list of cell states, characters can be identified. This is shown for a 7-segment display in **Figure 6**. Similar lists can be developed for other field types. Note that the term character is used rather loosely and does apply to elements formed from display symbols.

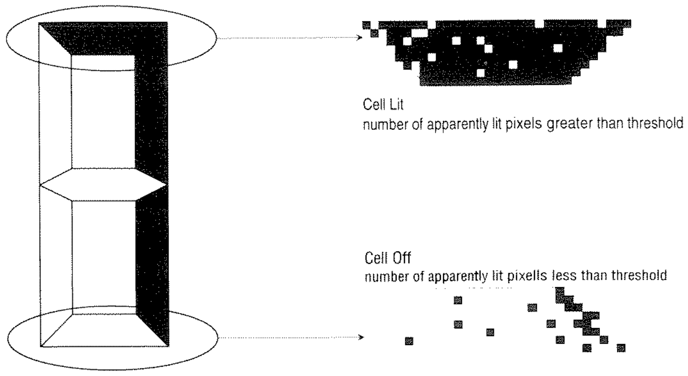

**Figure 5.**
Determination of cell illumination.

Note also that this algorithm works for dot matrix characters where the position of the characters is fixed on the display; however, where characters are of variable size, or can start at any point in the display, other algorithms must be used in conjunction with the one described here.

For each field, the character displayed by each element can be determined from the cell list. Dependent on the type of information held in the field the spoken phrase can now be constructed. A simple example of an alarm clock is shown in **Figure 7**. The alarm clock display consists of three fields: 1) a time field made up of four 7-segment elements; 2) an a.m./p.m. field which is a single static element (this is lit to indicate "p.m."); and, 3) an alarm set field which is a single static element (this is lit to indicate that the alarm is set). To speak the time, the display reader must chain together the phrases associated with the time and the a.m./p.m. fields. The alarm set field is spoken independently and is not considered further here.

The phrase for the time field is held as: "The time is ~." The character ~ indicates that the contents of the field should be spoken at this point. Thus, the field contents can be inserted into any position in an arbitrary text string. The field is marked in the configuration store as being time format; thus, the set of characters 1 2 3 4, is spoken as "twelve thirty-four," rather than "one thousand two hundred and thirty-four," which would be the case if the field was marked as a value field. Field 2 (a.m./p.m.) is marked as an alternative field, where two alternative phrases are stored, the phrase used being dependent on the state of the field: "pee em./aye em."

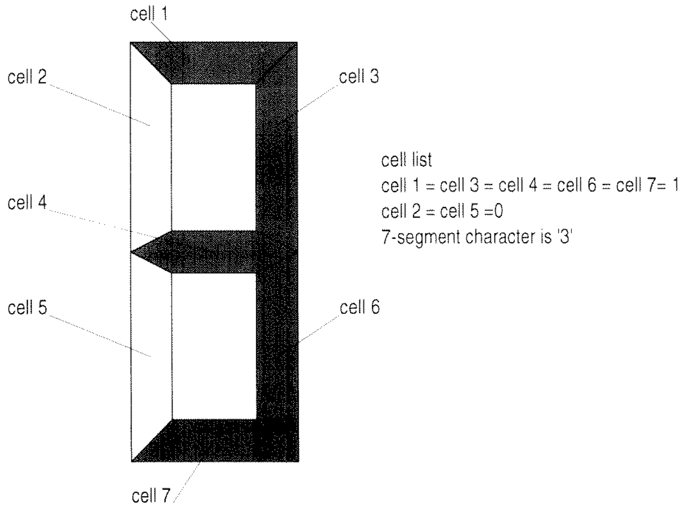

**Figure 6.**
Determination of cell states for a 7-segment display character.

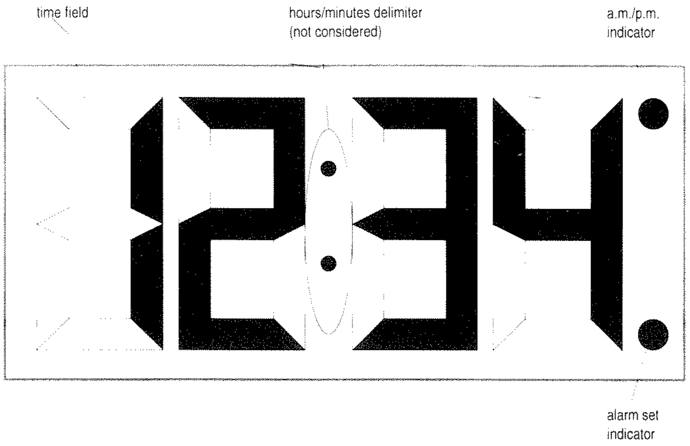

**Figure 7.**
An alarm clock display.

Dependent on the current operating mode of the system (see User Interface section above), which is held in the Display Configuration Store, and on the values obtained from the display, the text strings may or may not be passed to the speech synthesizer unit. Considering the alarm clock example, if the mode was set to read field on change, the reader system samples the display about once per second and compares the value of the time field with the value captured previously. Text is spoken when a change is detected; in this example text would be spoken every minute. If the user requested that a field be spoken, the field would be returned and spoken immediately.

## RESULTS

### Prototype Evaluation

*Display Quality*

The quality of the captured image has been evaluated by capturing images from the displays of a number of domestic products, which cover the range of available display technology (i.e., LEDs, GPDs, and LCDs) both backlit and lit by ambient light. The quality has been assessed by visual inspection following the thresholding of the image; in all cases lit cells were clearly distinguished.

*System Evaluation*

The prototype system has been evaluated by operating the system with three domestic products. Again, a sample representative of the range of image technologies was chosen; namely LED, GPD, and ambient lit LCD. The system has also been evaluated by running simulations of product displays on portable computer systems. This allows the emulation of complex displays, yet gives total control and repeatability to each of the experiments. Most of the work to date has been carried out using an emulation of a video recorder, which has representatives of all the elements found in the complete set of domestic products. This emulator has been run on two PCs; one with a GPD and the other with backlit LCD. This testing has verified that the image processing algorithm is robust and virtually error free over a reasonable range of display technologies. One small error has manifested itself, however. When reading displays in *read field on change* mode the display image is sampled once every second, for 7-segment displays the character "0" is read in error, on average once every 20 minutes, as the character "8." This manifests itself by the system indicating that the channel of the emulated video recorder display has changed from "channel 2" to "channel 82" and on the next read back to "channel 2." This is equivalent to an error rate of approximately 0.08 percent. Only 7-segment zeros seem to suffer from this defect.

### Image Alignment

As noted above, the image processing algorithm requires that the camera and display be in the same relative alignment as when the system was configured. System configuration takes a significant period of time, around 20 minutes for a video recorder display. Any movement of the camera relative to the display, even if this movement is comparatively small, results in the system having to be reconfigured.

As noted earlier, such precise alignment will be difficult, if not impossible, to maintain for a practical system. Therefore, work will be undertaken to permit the system to align a captured image with the information held in the configuration store.

### Fishbowling of the Image

The image alignment problem is compounded by the fact that the image passed by the Peach camera to the system is fishbowled (i.e., straight segments at the edge of the image appear curved). This is because the Peach camera has a wide-angle lens, and in this application the camera is in rather close proximity to the display. This problem could be overcome by either obtaining a miniature close-up lens for the Peach camera, or by applying some sort of a corrective transformation to the captured image.

### System Configuration Policy

As has been discussed above, the system will need a specific configuration for each appliance with a display. It is envisaged that, although facilities will exist for a sighted user to produce configurations for a given individual, the configurations generally will be produced and distributed by a central agency, probably the supplier of the display reader. Although this will involve a considerable effort initially, it is expected that subsequent efforts will only involve keeping up to date with the latest product developments. There are issues relating to upgrading individuals with configurations for the latest devices; however, this is well within current capabilities for electronic data transfer and is not dealt with here.

## DISCUSSION

The ultimate aim of this work is to produce a handheld product, at a reasonable cost, that is suitable for use in the home. To this end four areas of further work are under consideration:

1. **Testing with a Wider Range of Products**. As outlined in the section concerned with prototype evaluation above, a limited yet representative subset of displays has been used to evaluate the system. More extensive testing is required and a systematic evaluation of the prototype system will be carried out with a larger set of displays.
2. **The Evaluation of the User Interface**. It is intended that the user interface will be evaluated by carrying out a series of experiments with visually disabled users of the system. The purpose of the work is twofold: to ensure that a suitable user interface is developed and to develop a more general level of understanding as to how visually disabled people can effectively operate with modern domestic appliances.
3. **Image Alignment**. Further work is necessary to address the image alignment problems and fishbowling discussed above. It is anticipated that this work will include experiments conducted in a similar manner to those in the previous section, with visually disabled users of the system.
4. **Development of the Target System**. Satisfactory completion of this work will result in a prototype of a highly integrated, hand-held product that can be used in the home. This will entail developing an ASIC-based system. This work will progress in par-

allel with other, related work undertaken by the systems engineering group at UMIST as part of the ESPRIT OMI/DE[1] project. It will result in the production of a deeply embedded processor-based system for a range of image processing applications, of which the system described here is one. The system will be based around an ARM[2] processor macrocell; this processor has high computational power and therefore is capable of meeting the real-time requirements of the display reader system and has low power requirements.

## CONCLUSIONS

A proof of concept demonstrator has been constructed and sufficient evaluation carried out to generate confidence that this approach to the problem of visually handicapped people accessing visual displays is valid. Despite its limitations, the prototype system can and should be used to evaluate the real needs of visually disabled users. As for subsequent development at the target price of $100 suggested by the World Blind Union Research Committee, this seems unlikely in the short term. However, improvements in silicon processing technology will allow a system to be integrated onto a single silicon chip in the future and would probably enable the system to be supplied at its target price in from 4 to 5 years from now.

## ACKNOWLEDGMENTS

This project was carried out by a group of five students as part of their Master of Engineering degrees in microelectronic systems engineering or in software engineering, at UMIST in the academic year 1992/93. Joe Carter, David Gent, Justin Peters, Chris Ross, and David Whipp developed the system from initial specification supplied by the Technology for Disabled People Unit (TDPU). In transforming this specification into the prototype system, these students have shown tremendous enthusiasm and innovative ability, and have put in a great deal of hard work. The authors thank them for their assistance. We would also like to acknowledge the major contribution to the work reported here, and to the TDPU in general, of our colleague Professor Derrick Morris. The TDPU would like to acknowledge research grant funding from the Guide Dogs for the Blind Association, which was used to fund this work.

## REFERENCES

1. Kay L. Electronic aids for blind persons: an interdisciplinary subject. IEEE Proceedings 1984:131(Part A):559–76.
2. Meyers A, Schreier E. An evaluation of speech access programs. J Visual Impairm Blindn 1990:84:26–38.
3. Schwerdtfeger RS. Making the GUI talk. BYTE 1991:December:118–28.
4. Boyd IH, Boyd WL, Vanderheiden GC. The graphical user interface: crisis, danger, and opportunity. J Visual Impairm Blindn 1990:84:496–502.
5. Gill JM, ed. Priorities for technical research and development for visually disabled persons. World Blind Union Research Committee, 1992.
6. Morris D, Evans DG. Modelling parallel and distributed systems. Parallel Comput 1992:18:793–806.
7. Blenkhorn P, Caulderwood D. Access to personal computers using speech synthesis: a review of the past decade. The New Beacon—Blind Welfare 1992:LXXVI(898):185–8.
8. Ronsse C. The IBM workstation for the blind. In: Francois G, Engelen J, eds. Proceedings of the 6th International Workshop on Computer Applications for the Visually Handicapped, Leuven, Belgium, 1990.
9. Talbot M, Blenkhorn P. Building context-sensitivity into special access products. In: Murphy H, ed. Proceedings of the Seventh Annual International Conference on Technology and Persons with Disabilities, 1992 March, Los Angeles, CA.
10. Vincent T, Turnbull SD. Word processing for blind people. Microprocessors Microsystems 1984:8:535–8.
11. Yourdon E. Modern structured analysis. Englewood Cliffs, NJ: Prentice-Hall, 1989.
12. Carter JA, Gent DP, Peters JL, Ross CC, Whipp DP. Display reader for the Blind. Microelectronic System Engineering and Software Engineering MEng., Project Report. Manchester, UK: University of Manchester Institute of Science and Technology, 1993.
13. Renshaw D, Denyer PB, Wang G, Lu M. ASIC image sensors. In: Proceedings of the IEEE International Symposium on Circuits and Systems, 1990.

---

[1] The European Community has a funding program for precompetitive research in the information technology industries to improve Europe's competitiveness. This is known as ESPRIT, part of which is concerned with the microprocessor/microelectronics sector and known as the Open Microprocessor Initiative (OMI). The Deeply Embedded(DE) project is investigating highly integrated systems that result in complete software/hardware systems on a single silicon chip.

[2] This refers to a family of high performance, low power consumption 32-bit RICS processors developed in the United Kingdom by ARM Ltd. The technology is licensed to a number of silicon vendors worldwide and the processors have been used in a large number of high-performance, battery-powered devices.