RESEARCH AND DEVELOPMENT IN THE FIELD OF READING MACHINES FOR THE BLIND

Glendon C. Smith, P.E.
Senior Project Engineer

Hans A. Mauch, Dipl. Ing.
President

Mauch Laboratories, Inc.
3035 Dryden Road
Dayton, Ohio 45439

ABSTRACT

Mauch Laboratories' Cognodictor reading machine for the blind has been redesigned to meet much higher standards of performance. Consequently, the new High Performance Cognodictor should satisfy the needs and expectations of a larger proportion of the veteran and non-veteran blind population.

Earlier Cognodictor designs had low cost as the dominant design criterion.

Significant progress has been made in the development of hardware and software for the new Cognodictor. A hand-held probe has been designed, built, and tested. Circuits which preprocess the video signal from the probe, and circuits which interface the probe with a PDP-11 minicomputer, have been built and placed in operation. Test and video display programs have been written and debugged. Routines in the recognition software which detect word spaces, locate character boundaries, and measure height and width of a character have been written and tested. More specific routines for the letters o, c, e, a, s, and t are being written and debugged.

It is planned to have the construction of a Cognodictor prototype parallel the completion of the software, using LSI-11 microcomputer hardware for character recognition, and 8080 microcomputer hardware with a Votrax synthesizer (which has been in operation at Mauch Laboratories) for speech production.

INTRODUCTION

The blind members of our society generally have low incomes and reading machines for the blind might be described as partial

replacements for relatively inexpensive systems of print access such as talking books, braille, and volunteer readers. In the past, Mauch Laboratories considered low cost as one of the most important design criteria for both direct-translation and recognition-type reading machines.

Earlier designs of Mauch Laboratories were considered quite successful in providing a good level of performance at low cost. For example, the Stereotoner, priced at $1,250.00, is half the cost of the Battelle Optophone and less than one-half the cost of an Optacon despite the much higher production level of the latter. Maximum reading rates of about 80-90 words per minute have been reported for both the Optacon and the Stereotoner, and average rates are comparable.

Utilizing only 130 low-cost integrated circuits, the first Cognodictor produced spelled speech for all the lowercase letters and most of the uppercase letters of about nine common styles of type. Considering its projected cost of about $4,000 in small quantities, its performance/cost ratio was considered excellent for its time.

The Two-Dimensional Multiple Snapshot Cognodictor extended the tolerance for mistracking by a factor of five and was designed using 19 type styles. Its output was to be spelled speech at rates up to 80-90 wpm (words per minute). Its cost was projected at $6,000. Before it could be completed, it became apparent that performance expectations had increased substantially and that cost was therefore much less a factor—at least, for the federal agencies and foundations which usually purchase and evaluate the first models of such devices. Accordingly, a new Cognodictor design was started in April 1976.

PROGRESS IN COGNODICTOR DEVELOPMENT

New Cognodictor Design

In April 1976, the sponsor approved changes suggested by Mauch Laboratories intended to lead to a reading machine with much higher performance—one which would have an error rate of less than one percent in letters and numerals of a wider variety of type styles. This High Performance Cognodictor was to have a synthetic speech output, with spelled speech or direct translation (tone code) outputs also available at the press of a button. The design called for two components—a high-resolution optical probe designed for hand-held use, and a control box containing two microcomputers and a Votrax vocal synthesizer. Initially the control box was expected to be about 2¼ ft^3 in volume and weigh 30-35 lb. The

Cognodictor's cost (in September 1977) is estimated at about $12,000 each in a quantity of 10. The price, size, and weight can all be expected to decrease substantially with the rapid progress apparent in electronics technology. Additional cost reduction should be possible with increasing production.

The designers believe that the combination of the hand-held probe of the Cognodictor and its direct translation output mode will combine to allow the user to understand the format of his document and decipher unusual symbols. In this mode, tonal or tactile patterns which correspond closely with the print will be produced. The vertical span covered by these patterns can be set by a user-operated magnification control, so that one line of print can be sensed at a time. The additional information conveyed in this mode would be difficult or impossible to obtain from an automatic page-scanning machine without a direct translation capability.

The Fairchild CCD-121 photocell array in the probe is designed to scan a band 1½ in high, thus providing a tracking tolerance of ±.675 in for typewriter-size print. This amount of tracking tolerance should make freehand tracking very fast and easy, eliminating the need for tracking aids in most cases. An automatic page-scanner, for use with long magazine articles and books, could be made available in the future.

The new Cognodictor design goals include a wider range of characters (including italics and numerals) over a wider range of type styles. The number of photocells used (~ 1500 from an array of 1728) permits both a large field of view and high resolution, which should yield reliable detection of lines as fine as .003 in (found in some common type fonts such as Caslon, Baskerville, Primer, etc.), even in the presence of interference from print on the other side. It can also allow adjustment for type size to be an automatic, electronic operation. (A manual override would be provided for use in direct translation mode when needed.) The photocells are sampled each time the probe moves .003 in horizontally as determined by an optical encoding disk which contacts a roller.

The design provides for automatic adjacent-line suppression based on information obtained by the photocells above and below the line being read. A similar process will be used for the automatic suppression of underlinings. This adjacent-line suppression can be overridden by a control on the main box, for the rendition (in the direct translation mode) of mathematical formulae, geometric figures, etc., covering more than one line-space.

Italics will not necessitate slanting the probe but will be automatically accommodated by the Cognodictor. The computer in the

Cognodictor will also be used to decide whether the print is typewritten (equally spaced) which will facilitate word separation (look for the middle of a letter location, if it is white it must be a word space)—or not typewritten in which case word and letter spaces are sufficiently different to avoid ambiguities in discovering word spaces. Touching letters will initiate a subroutine to separate and recognize them.

It is intended to have the Cognodictor's direct translation mode initially use 10 tones, as in the Stereotoner, but monaural. The tones will be initiated by pressing a button on the probe. Later, an accessory tactile display could be developed, possibly one which embosses a reusable plastic sheet or strip.

The usual output, synthetic speech, will be produced by a vocal synthesizer such as the Votrax unit. Its output rate will be adjustable up to 150 words per minute. An additional computer program will allow the synthesizer to produce spelled speech at up to about 40 wpm. The basic pitch of the voice would be adjustable by the user to his preference; changes above and below the basic pitch are machine-controlled.

The design calls for the speaking of a word to begin as soon as possible after the detection of a word-space, to eliminate delays which would be present if the program were to require the user to scan to the end of the line or to the end of the sentence. For this reason, and to keep the size and cost of the machine within reason, it is not intended to use syntactical analysis and stress assignments in the Cognodictor at the start. Punctuation marks will be recognized and their names pronounced only when the spelled speech mode is used.

Probe Design

From previous exposure to the optical probes or cameras of direct translation reading aids such as the Stereotoner or the Optacon, one might assume that a comparable probe would meet the letter recognition needs of the High Performance Cognodictor being developed. Actually, the Cognodictor requires a probe with much greater capability. The greater accuracy and higher reading rates of the Cognodictor require a faster photocell array with more resolution than slower direct-translation devices require or can utilize. And, at higher scanning speeds, the user cannot track as accurately, so the vertical extent of the array must be increased. Therefore, the new probe (Fig. 1) has many more photocells (1728 as compared to 144 in the Optacon), the field of view is greater (1.5 inches as compared to .750 in the Stereotoner and .225 in the Optacon), and the speed of response of the photocells is much faster (0.5 ms

as compared to 5.0 ms for the Optacon). The probe also contains an optical encoder—not required in the Stereotoner or the Optacon—which informs the Cognodictor of the distance and direction of each horizontal motion.

It might also appear that using a hand-held probe simplifies the Cognodictor design as compared with using an automatic page-scanner mechanism. Actually, the Cognodictor/probe combination

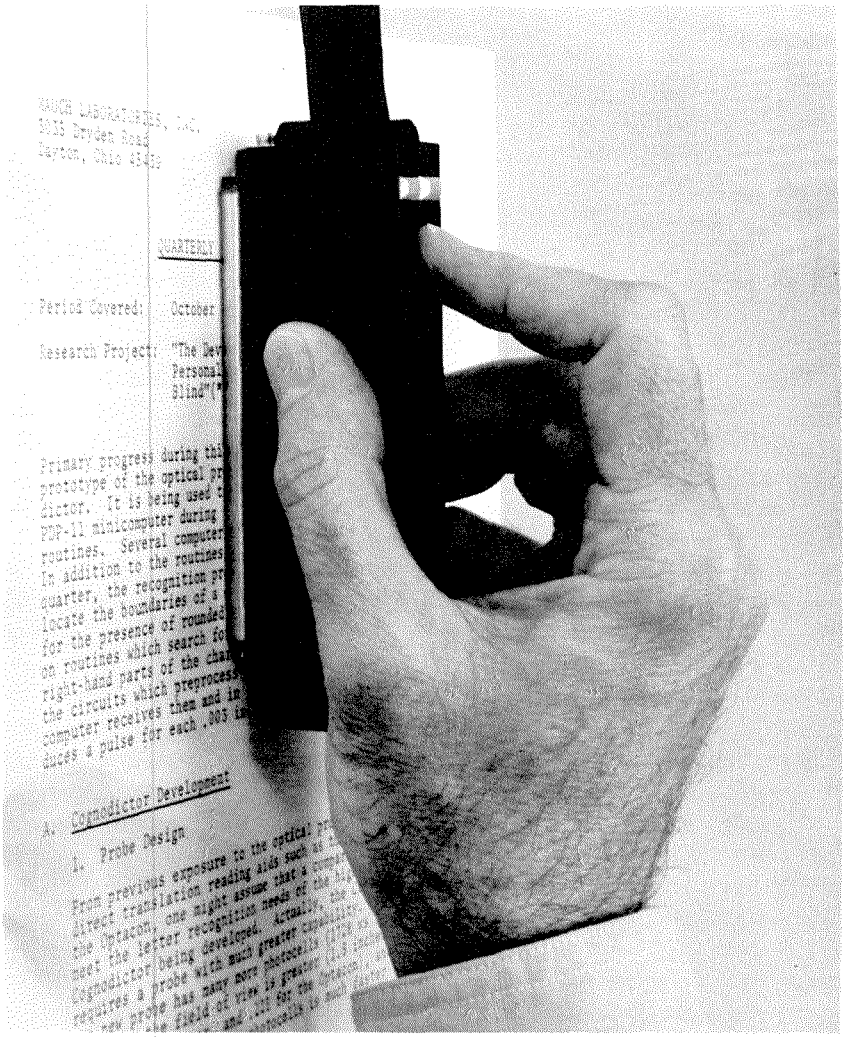

FIGURE 1.—The large field of view (1.5 in high) and long high-friction rollers of the Cognodictor probe make freehand scanning of print easier than ever.

is more difficult to design. Some of the reasons for this are given in the following comparisons between possible designs for probes and page scanners.

The illumination of the probe is limited to below about 2 W by the problems of disposing of waste heat, whereas the page scanner can use 20-50 W or more. In the page scanner uniform illumination can be achieved easily by diffusion and/or by moving the lamps several inches from the print: the probe must use efficient reflectors in a limited space.

The space available in a page scanner allows the use of an off-the-shelf lens designed for making close-up photographs with a 35 mm camera. The high level of illumination present allows the use of a small-aperture setting, and filters to remove infrared wavelengths, for a sharper image on the photocells. The space is also adequate to mount the photocell array directly to the circuit board that processes the signal, thus avoiding the electrical effects of a cable at a critical point.

The hardware and software designed for use with a probe must be able to handle a wide range of scanning speeds, while the page scanner may only have to function at one speed. For example, suppose both machines are to produce 150 wpm of synthetic speech and both are to scan typewritten pages with 10 characters per inch. To achieve this net rate, the probe user is expected to scan the typical 6-in line left-to-right in about 2 s, flyback and change to the next line in the next 2 s. This requires that he scan in the forward direction at 300 wpm or about 3 ips (inches per second). Actually, he cannot hold this or any other speed with precision, so the probe system should be designed for operation up to 6 ips. By comparison, the page scanner could scan and recognize in both directions at 2 ips and still produce 150 wpm with a 30 percent allowance for line change.

The designer of the page scanner also can be certain that each scan will be along a straight line with a preselected slope, so few photocells have to be allocated for possible mistracking, and the recognition program can be less complex than one which has to search over a larger area for a character feature.

Although these factors, among others, make the page-reader easier to design and build than the hand-held probe, the presence of an unlimited number of print formats among the materials the blind wish to read makes the probe concept worthwhile. It is expected that, after some training, the probe user will be readily able to scan new formats in reading matter, to locate short items in reference or text books, and to make sense of unusual indicia which are not recognized correctly by the machine.

The Cognodictor will have a page-scanner as an accessory, but with its hardware and software designed to meet the more stringent needs of a probe, adding a scanner should be relatively simple. Then, Cognodictor users will have the advantages of both systems.

Probe Construction

An optical encoder, which provides forward/reverse signals and clock pulses for each .003 in the probe rolls horizontally, has been designed and built. The success of the probe is very dependent upon the accurate and reliable operation of the encoder. Three months of continuous operation were used to test the durability of the encoder and associated circuitry.

Inside the encoder, a rotating transparent disk with 135 equal opaque sectors interrupts two beams of infrared light from two light-emitting diodes to two phototransistors. The interruptions produce square wave signals from the phototransistors which are equal except for being separated by a time displacement of one-quarter of the wave period (90 deg). A circuit detects which wave is leading the other in time and thus determines the direction of travel of the probe. Short pulses are produced by each edge of each waveform to mark the passage of each increment of .003 in.

The probe uses a reflex optical system with a 45 deg front surface mirror, a 20 mm focal length lens (later to be 18 mm) and a Fairchild CCD-121 photocell array. The magnification produced is 0.5×, a reduction of 2 to 1. In addition to the optical encoder and associated circuits, the probe contains two miniature switches and a circuit board containing an amplifier and several circuits needed to operate the photocell array. A flat cable about 1 in wide contains 20 wires which connect the probe to the Cognodictor circuitry. The overall dimensions of the probe are 5 × 0.8 × 1.7 in.

Preprocessing and Interface Circuits

To provide ample tolerance for mistracking (±.675 in for typewriter size) and a 5:1 range of acceptable print sizes, the Cognodictor probe has a field of view of 1.5 in. About 1,500 photocells sense this space so the vertical resolution is nearly as much as the lens quality will allow. Horizontal resolution is about 333 samples/inch.

The preprocessing circuits utilize two numbers (a tracking number and a lumping number) produced by either the user or the computer to determine which photocells are over the desired line and to combine samples from these cells to yield a standard height for the electronic letter image transmitted to the computer. The horizontal

samples are combined in the same proportion (given by the lumping number) as the vertical samples (photocells) to yield a width which is appropriate for the letter height. When the combining process has been completed, the interface circuits interrupt the computer and the computer accepts four 16-bit words (64 bits which comprise one column of points in its internal letter image). These preprocessing and interface circuits were assembled on two plug-in circuit boards each 4½ X 6½ in. They use 47 integrated circuits of the low power Schottky TTL family.

Cognodictor Software Development

During July 1976, a systems analyst, Dr. Chung C. Lee, began writing and testing computer routines which will eventually be combined into a large program which simulates almost all of the Cognodictor's functions. When sufficiently perfected, this program will form the basis for the Cognodictor software. Also, in July, a test input-output routine was written. In August, the development of a video display program was started. This display program causes the cathode ray tube to display the contents of the input buffer area of the computer's memory consisting of 1,024 words, of 16 bits each, which form images of the last 10-12 characters scanned by the probe. (Before the two interface boards between the photocell development board and the PDP-11 computer were operating, this buffer area contained synthetically produced letter shapes which were entered by hand.) The program also draws five horizontal lines delineating the boundaries between the four 16-bit words which comprise each column of points in the display.

The display routine, an initialization routine and an interface input routine were completed and debugged during September and October. Test signals were provided by the photocell array which scanned print attached to a constantly rotating drum. The interface circuits were operating except for the substitution of a fixed frequency clock pulse for the optical encoder signal.

In November 1976, three computer routines were written. The first one locates the left-hand and right-hand edges of the next character to be recognized and also yields a measure of its width. The second routine then locates the upper and lower edges and determines character height. Considering the rectangle whose sides coincide with these four edges as the character space, the third routine checks each inside corner of that space for black. The results of checking four corners yield 16 combinations of results.

During December 1976, the recognition routines which had been written to date were tested, using the hand-held probe to scan several hundred lowercase letters. The printout produced indicated

that the routines operated as intended. In January 1977 the next group of routines was completed. These routines will detect a vertical or nearly vertical straight line near the left or near the right edge of the character and will measure the length of each such line found. During February 1977, a number of routines needed for specific tests in the recognition of lower case o, c, e, a, s, and t were developed.

FUTURE WORK

Prototype Construction

Still to be assembled are the two circuit boards needed for the speech synthesis portion of the prototype. It appears that the printed circuit board which was designed to hold 16 programmable read-only memory circuits (PROMs) for the previous Cognodictor can be used for the PROMs which will contain the speech synthesis program. A new circuit board will be needed for the 8080 microprocessor, read-write memory, input-output ports, and associated circuits. Perhaps a suitable board can be purchased; if not, designing one should not be difficult.

Dr. Scott Allen at NIH has been writing a new version of the speech synthesis program we have been using. If the new version is available it can be loaded into the PROMs: if not, the older program can be used until a better one is available; at that time, the PROMs can be erased and reprogramed. The compact speech synthesis system consists of two circuit boards (each 6.5 in × 4.5 in) and a Votrax VS-6 (11.5 × 11.5 × 3 in).

It will be necessary to purchase the LSI-11 components for the recognition portion of the prototype. Thus far, a readymade plug-compatible PROM board for the LSI-11 has not been found and it may be necessary to prepare artwork and have a local circuit board manufacturer make several such boards (existing commercial PROM boards use fusible link PROMs which cannot be reused).

Software Development

The design of the optical probe, and the choice of an LSI-11 microcomputer for the recognition portion of the Cognodictor, provide a basis for performance which will be determined mainly by software. At present, the software capabilities are limited mainly by the time and manpower available for its design.

The first version of the recognition routines includes only the basic routines for recognizing uppercase and lowercase letters (excluding italics). Punctuation and numeral recognition are to be

added in Version 2. Separation of touching characters would be a main feature of Version 3. Automatic line-following and automatic adjustment of type size are probable additions to Version 4. Later versions can add routines for recognition of italic fonts, and routines which operate an automatic page-scanner. Each version should also contain modifications which improve earlier routines, as user experience reveals the need for such improvements.

Prototype Evaluation

It is expected that Mr. Harvey Lauer at VA Hospital, Hines, Illinois, will be the first outside user of the High Performance Cognodictor prototype. Mr. Lauer is already skilled in probe manipulation and in using the aural direct-translation output code. It is hoped that previously unskilled potential users of the Cognodictor, such as veterans being rehabilitated at Hines, will receive exposure to the Cognodictor and some training in its use and that the feedback from these individuals will be available to the designers. After about 8 months the prototype could be moved to another situation (such as a business office, school, or home) for additional testing.

Other Developments

When the effort required to maintain and improve the Cognodictor software and hardware diminishes, work could begin on the design of an automatic page scanner accessory. Both the basic design of the device and its interface with the operator require thorough study. For example, the operator controls might be either a specialized keyboard with keys for different functions, or a standard full keyboard as on a typewriter (with single letters or mnemonic combinations of letters as the control codes) or a "joystick" controller with a few auxiliary pushbuttons and controls.

Perhaps the choice should be influenced by possible future extensions of the Cognodictor's capabilities. Some of these possibilities are:

1. With the standard keyboard and additional software the Cognodictor could become a talking intelligent remote terminal for a bigger computer, or it could be a stand-alone minicomputer if a mass storage device such as a floppy disk unit is added.

2. With the above peripherals and a hard copy output such as a Selectric printer, the performance of a magnetic card typewriter could be duplicated. Other input-output combinations are possible, utilizing braille line printers, paper tape readers, or telephone line connections.

3. The speech generation components (8080 microcomputer and Votrax) could be used separately as part of a Teletype replacement for blind programers or workers such as reservation clerks, who interact with a computer.

In conclusion, it may be a waste of an expensive prosthetic appliance to utilize the power of the microcomputers and the voice synthesizer in the Cognodictor only for reading—if other functions can be added as options at reasonable costs. This possibility should be studied in detail.