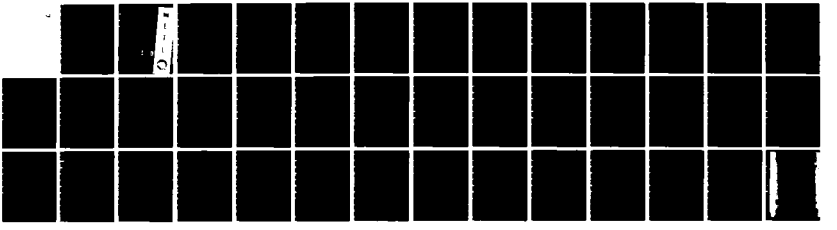
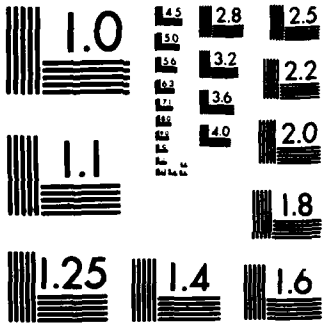
AD-A142 698 VOICE INTERACTIVE SYSTEMS TECHNOLOGY (VIST) RESEARCH (U) ARMY ENGINEER TOPOGRAPHIC LABS FORT BELVOIR VA T F DEYOUNG JAN 84 ETL-0349 1/1

UNCLASSIFIED F/G 9/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A142 698

ETL-0349

Voice interactive systems technology (VIST) research

Tice DeYoung

JANUARY 1984

DTIC FILE COPY

DTIC
ELECTE
JUL 0 6 1984
S D
E

U.S. ARMY CORPS OF ENGINEERS
ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA 22060

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

84 07 05 088

E T L

U.S. ARMY ENGINEER TOPOGRAPHIC LABORATORIES
ETL

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

The citation in this report of trade names of commercially available products does not constitute official endorsement or approval of the use of such products.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: ETL-0349	2. GOVT ACCESSION NO.: AD-A142698
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): VOICE INTERACTIVE SYSTEMS TECHNOLOGY (VIST) RESEARCH	5. TYPE OF REPORT & PERIOD COVERED: Interim Technical Report October 1981–September 1982
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Tice F. DeYoung	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia 22060	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 6.27.07.A 4A 762707A855 C 0031
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia 22060	12. REPORT DATE: January 1984
	13. NUMBER OF PAGES: 38
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Automated Word Recognition
Speech Recognition
Speech Synthesis
Voice Technology
VTAG

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report discusses the results to date of the VIST research at the U.S. Army Engineer Topographic Laboratories (ETL) and describes in detail how to make the Interstate Electronics Corporation Voice-Recognition Module and Voice-Synthesis Module work on Hewlett-Packard 1000, Data General Eclipse S-250, and Digital Equipment Corporation PDP-11/45 minicomputers.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

PREFACE

This research was conducted under DA Project 4A762707A855, Task C, Work Unit 00031, "Voice Interactive Systems Technology Research."

The research was done during the period October 1981 to September 1982 under the supervision of R.D. Leighty, Team Leader, Center for Artificial Intelligence; and M. Crowell, Jr., and R.D. Leighty, Directors, Research Institute.

COL Edward K. Wintz, CE, was Commander and Director and Mr. Walter E. Boge was Technical Director of the U.S. Army Engineer Topographic Laboratories during the report preparation.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC COPY INSPECTED 2

TABLE OF CONTENTS

SECTION	TITLE	PAGE
I.	INTRODUCTION	4
II.	BACKGROUND	5
III.	INSTALLING THE VIST HARDWARE	7
	A. Hewlett-Packard 1000	7
	B. Data General Eclipse S-250	10
	C. Digital Equipment Corporation PDP-11/45	11
IV.	TESTING THE VIST HARDWARE	15
	A. VRM Test Programs	15
	1. HP-1000	15
	2. DG Eclipse S-250	17
	3. DEC PDP-11/45	19
	B. VOTALK Test Programs	20
	1. HP-1000	21
	2. DG Eclipse S-250	21
V.	DEMONSTRATING VIST CAPABILITIES	22
	A. HP-1000	22
	1. VOTERM Functional Subroutines	22
	2. Image Processing Demonstration Program	25
	3. DEMONS Demonstration Program	27
	B. DC Eclipse S-250	30
	1. VRM Functional Subroutines	31
	2. Vertical Obstruction (VO) Demonstration Program	31
	C. DEC PDP-11/45	33
	1. VOTERM Functional Subroutine	33
	2. Graphics Manipulation Program	34
VI.	RESULTS AND CONCLUSIONS	36
VII.	REFERENCES	38

TABLES

NUMBER	TITLE	PAGE
1	VRM REQUIREMENTS FOR PROPER OPERATION WITH HP-1000 MINICOMPUTER SYSTEM RUNNING UNDER RTE-IVB	8
2	VRM REQUIREMENTS FOR PROPER OPERATION WITH DG ECLIPSE S-250 MINICOMPUTER RUNNING UNDER AOS	12
3	VRM REQIREMENTS FOR PROPER OPERATION WITH DEC PDP-11/45 MINICOMPUTER RUNNING UNDER RSX-11M	13
4	VOTERM RECOGNITION VOCABULARY FOR IMAGE PROCESSING VIST DEMONSTRATION PROGRAM ON HP-1000 MINICOMPUTER	26
5	VOTERM RECOGNITION VOCABULARY FOR DEMONS VIST DEMONSTRATION PROGRAM ON HP-1000 MINICOMPUTER	28
6	VRM RECOGNITION VOCABULARY FOR VERTICAL OBSTRUCTION DEMONSTRATION PROGRAM ON DG ECLIPSE S-250 MINICOMPUTER	32
7	VOTERM RECOGNITION VOCABULARY FOR GRAPHICS MANIPULATION PROGRAM ON DEC PDP-11/45 MINICOMPUTER	35

VOICE INTERACTIVE SYSTEMS TECHNOLOGY (VIST) RESEARCH

I. INTRODUCTION

The Voice Interactive Systems Technology (VIST) Research at the U.S. Army Engineer Topographic Laboratories (ETL) considers the application of existing voice technology to the problem of efficient and accurate extraction of cartographic, topographic, and military intelligence information from aerial imagery. The VIST research at ETL was initiated to investigate currently available VIST devices and to determine where they could be most effectively used in systems under development at ETL. In most of these systems, the operator and decision-maker is the expert photo-interpreter, analyzing high-resolution, stereoscopic aerial photography. Although great progress has been made in stereoscopic optical viewing systems and in the development of computer assisted devices such as the ETL's Computer-Assisted Photo-Interpretation Research (CAPIR) system, a great deal of work still remains in order to improve the interaction between the human analyst and the system.[1]

For example, while using the CAPIR, the operator periodically has to shift his/her eyes from the stereomodel to the CRT and remove his/her hands from the trackball and thumbwheel switch to feed data into the system via the keyboard. This detachment from the primary task slows the overall data input rate and can even cause some errors in interpreting the image. VIST can provide a means to eliminate this undesirable eye and hand movement between the stereomodel and the CRT by supplementing the prompts on the CRT with audible prompts through an earphone, using a synthetic speech device, and by allowing the operator to input data and commands with a speech-recognition device instead of typing the information in via the keyboard.

---

[1] G.E. Lukes, "Computer-Assisted Photo-Interpretation Research (CAPIR): A Prospectus," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA, 22060, November 1979, unpublished

VIST devices have been shown to be extremely useful when an operator has his/her hands and/or eyes busy performing tasks like image interpretation, since they transfer some of the tasks from the hands and/or eyes to the virtually unused speaking and hearing capabilities of the operator. VIST devices can also be of great utility in image display and manipulation, again freeing the operator's hands and eyes to perform the primary task.

The first section of this report describes how to install and test VIST hardware on three different minicomputer systems used for research in the area of computer-assisted image interpretation at ETL: A Hewlett-Packard 1000, a Data General Eclipse S-250, and a Digital Equipment Corporation (DEC) PDP-11/45. The remainder of the report is organized as follows: Section II gives some of the background leading to the current research. Section III deals with installing the VIST devices. Section IV discusses testing the VIST hardware. Section V gives the demonstration programs for each of the three minicomputers. The results and conclusions to date are presented in section VI.

II. BACKGROUND

An initial effort into VIST was begun in FY80 at ETL as an In-House Laboratory Independent Research (ILIR) project. The purpose of the ILIR was to look at VIST, using low-cost equipment, to determine if the technology could be used by any of the systems under development at ETL. This research used an IMSAI microcomputer, a Heuristics Speechlab voice recognition unit, and a Computalker speech synthesizer. With this equipment, several demonstration programs were developed that would move a cursor on a CRT screen, display selected images from a magnetic tape, and manipulate the images on a display monitor. The results of the ILIR project showed that even with very crude, low-cost equipment, VIST could be useful. Systems noted at that time as possible candidates for applications of VIST were the Computer-Assisted Photo-Interpretation (CAPIR) system and the Demonstration System (DEMONS).

As a result of the ILIR findings, it was decided to initiate a 6.2 work unit in VIST. This effort used the experience gained in the ILIR to determine what could be accomplished with more moderately priced hardware.

In FY81 and the first part of FY82 several pieces of VIST hardware were purchased. The voice-recognition hardware consisted of an Interstate Electronics Voice-Recognition Module (VRM) Model VRM-102 in a chassis (VOTERM). The VRM-102 is a speaker-dependent, isolated word recognizer with a 100-word-recognition vocabulary. Several VRMs were purchased in the VOTERM chassis, and one VRM was purchased without the VOTERM. This VRM was installed in the electronics section of the APPS-IV computer-assisted stereoscope on the CAPIR system. One VRM was purchased installed inside a Lear-Siegler ADM-5 CRT, for testing voice and data input from a single work station. Two different voice synthesis units were purchased, a Federal Screw Works VOTRAX and an Interstate ELectronics VOTALK. Both of these voice-synthesis devices are phonemic synthesizers, that is, they use phonemes for creating the speech. This type of device has the advantage of an unlimited vocabulary, but this is at the cost of a slight loss in intelligibility.

The VRM and VOTALK were installed and tested on a Hewlett-Packard (HP) 1000 series minicomputer. Several demonstration programs were written that would display and manipulate digitally stored images using voice commands and audible prompts, and data were input and verified using the VIST hardware. These successful demonstrations led to the decision to install the VRM and VOTALK on a Data General (DG) Eclipse S-250 minicomputer and on a Digital Equipment Corporation (DEC) PDP-11/45. These two minicomputers are the same as, or similar to, those used for the CAPIR and DEMONS systems, respectively.

III. INSTALLING THE VIST HARDWARE

Installing equipment made by a minicomputer company on its own system is relatively simple. However, installing anything that is not manufactured by that company can sometimes be a complicated and frustrating experience. Proper signals must be present for a correct serial (RS232C) or parallel connection to be made. In addition, certain communication protocols must be followed and these protocols vary with the manufacturer. This section of the report discusses which RS232 signals are required and the proper communication protocols needed to install VIST hardware on the HP-1000, DG Eclipse S-250 and DEC PDP-11/45 minicomputer systems at ETL.

A. Hewlett-Packard 1000

The Hewlett-Packard (HP) 1000 minicomputer in the ETL Research Institute's Center for Artificial Intelligence runs under HP's RTE-IVB operating system. Any device used on this system has to have an interface card installed in the system and an assembly language driver routine for proper operation. The RTE-IVB operating system must be generated using an HP program called RTGEN. This program uses a cross-reference table for location within the system and for equipment type and driver routine for each device installed.

The VRM board in the VOTERM has switches that must be set to accommodate the various communication protocols needed by the host minicomputer system. The settings for the HP-1000 are given in table 1 along with the necessary RS232 connections and the system generation information.

There are two RS232C serial interface cards for the HP-1000: the 12531D and the 12966A. The first one will not support any non-HP devices without occasionally losing data. The second one will support non-HP equipment, but there is no HP driver routine available for nonstandard HP equipment. A nonstandard driver was obtained from NASA for the VOTERM. This driver, called DVB00, worked successfully for both the VOTERM and VOTALK devices and makes the 12966A interface board completely programmable under software control, thus allowing flexibility in its use. This flexibility, however, puts a great deal of responsibility on the user; for instance, each program that will access the 12966A card has first to initialize it.

TABLE 1. - VMR requirements for proper operation with HP-1000 minicomputer system running under RTE-IVB

CABLE CONNECTIONS

Pin 1 - Ground
Pin 2 - Data out of Computer to VRM
Pin 3 - Data into Computer from VRM
Pin 7 - Ground

HP-1000 SYSTEM GENERATION INFORMATION

RTE-IVB Operating System Rev. 2126
Uses system generation program RTGEN
Use nonstandard HP Driver, DVB00 from NASA
Use 12966A Interface Board with Option 60004 Cable
(N.B. Need to reverse data lines 2 & 3 on cable)

VRM SWITCH SETTINGS FOR HP-1000

Switch SA: Set for 2400 baud (SA3 closed)
Note: If auxiliary display is used, VRM must be set for for 1200 baud (SA4 closed)
(N.B. All others on SA must be left open)

Switch SB: SB3: Open
SB4: Closed
SB5: Open
(All others on SB are don't care)

Switch SC: SC5: Closed
SC6: Closed
SC7: Closed
SC8: Open
(All others on SC are don't care)

Switch SD: RS232 I/O for Port 1 (Switch Label C1)
RS232 I/O for Port 2 (Switch Label C3)

Switch SE: Bypass preamplifier (Switch label - OUT)
RS232 logic levels for Port 2 (Label - PORT 2)
(N.B. For VRM not in VOn VOTERM II, set switch to Preamp IN)

The initialization begins by issuing a master reset command to the 12966A card. This master reset command clears the interface buffer and sets all the control lines to the proper state required by the VOTERM and VOTALK. The master reset command also sets the data transmission rate to 300 baud, thus requiring a second command to set the baud rate to the desired speed of 1200 baud for the VOTERM or 2400 baud for the VOTALK. Care must be taken when changing the baud rate to keep the control lines in the proper states.

Once the initialization of the 12966A board has been completed, communications between the HP and the VIST hardware may be performed using the normal HP formatted READ and WRITE statements. However, the user must again use caution to be sure that the commands will be recognized by the VIST hardware, because the format for the commands differs between the VOTERM and VOTALK. It is necessary to define the commands for the VOTERM first before attempting to use it. The HP control commands for the VOTERM must have the ASCII control character DC1 in the high order byte (upper 8 bits) with the low order byte (last 8 bits) being used for the ASCII equivalent of the desired function. The VOTERM's response to the command has the ASCII control character DC2 in the high order byte and its ASCII single character responses in the low order byte. The commands for the VOTALK and its responses are all standard ASCII characters and do not need to be predefined. The responses from both systems can be checked against those that are expected for proper execution of the the desired command. A more detailed explanation of how to use the VOTERM with the NASA driver, DVB00, on the HP-1000 minicomputer is found in other reports.[2] [3] [4]

[2]T.F. DeYoung, "Preliminary Speech Recognition/Synthesis Experiments," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060, September 1980, unpublished report.

[3]A.R. Wildes, DVB00, A General Purpose, Multiple Device Driver for the HP 12966A Buffered Asynchronous Date Communications Interface, NASA, Goddard Space Flight Center, Greenbelt, MD 20771, March 1977.

[4]T.F. DeYoung, "Communication Protocols for Making the Voice Recognition Module (VRM) 'Talk' to the HP-1000 Minicomputer System," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060, report in preparation.

The programs written to test the VIST hardware on the HP-1000, and those written to demonstrate what can be done with VIST are discussed in sections IV and V.

B. Data General Eclipse S-250

The Data General (DG) Eclipse S-250 minicomputer in the Research Institute's Center for Artificial Intelligence (CAI) runs under the AOS operating system. Any device used on this system has to have its requirements specified when it is generated into the system via the system console. The Eclipse S-250 uses the ALM Asynchronous Communication System for its peripherals. The VOTERM appears to the system to be a standard input/output device.

The VRM board on this system is installed in the electronics section of the APPS-IV. The VRM does not interact with the APPS-IV, but does get its power from the same power supply. The VRM board switch settings for the Eclipse S-250 communication protocols are presented in table 2, as are the RS232C connections and the system generation information.

The Eclipse S-250 has standard RS232 ports available for the VOTERM. However, the system must be in a HALT mode before any connections are made or broken. If the system is not halted, making or breaking these connections may cause the system to fail.

Since the VRM acts to the Eclipse S-250 like a standard RS232 device, there are no system initialization commands. However, every program that uses the VRM must let the system know where the VRM is located by using the FORTRAN callable routine OPEN. The VRM commands for the Eclipse S-250 must be predefined because its operating system, AOS, like that for the HP-1000, does not properly handle control characters. Additionally, AOS uses a line-feed (LF) as the terminator for its standard FORTRAN READ and WRITE functions, while the VRM requires a carriage-return (CR) or CR/LF as the terminating character(s). Therefore, it was necessary to use arrays for the VRM commands, rather than the single word commands used on the HP-1000 system. The high order byte (upper 8 bits) of the first array element must have the ASCII equivalent of the control character DC1, while the low order byte has the ASCII equivalent for the desired function on the VRM. Depending on the desired command, one or more additional array elements are required. Regardless of the number of array elements, the final element has to contain a terminator recognized by the VRM. For our purposes we chose the CR/LF. The AOS FORTRAN callable routines

RDSEQ and WRSEQ were used for reading from and writing to the VRM, respectively. These functions require the number of bytes to be passed as one of their parameters. Responses from the VRM will again contain the ASCII control character DC2 in the high order byte with a single ASCII character response in the low order byte; however, as for the commands, this is only the first element of an array. The number of elements in the array depends on the command given to the VRM, but here again the final element is the terminator, CR/LF. Proper response from the VRM to the commanded function can again be checked by comparing it to the response for correct operation. A more detailed explanation of how to use the VRM with the AOS operating system on the DG Eclipse S-250 minicomputer is given in another report that is in preparation.[5]

The VOTALK voice synthesis system has not been extensively used on the Eclipse S-250 System, so it has not been included in the discussion for this minicomputer. It should be noted, however, that plans for the CAPIR system do include using a speech synthesis device like the VOTALK in the future, and the VOTALK may be installed and tested on this system in FY83.

The programs written to test the VRM on the Eclipse S-250 and those written to demonstrate what can be done with it on this system are discussed in sections IV and V.

C. Digital Equipment Corporation PDP-11/45

The Digital Equipment Corporation (DEC) PDP-11/45 minicomputer in the ETL Topographic Development Laboratory's Automated Cartography Branch runs under DEC's RSX-11M operating system. A VOTERM is connected to one of the system's RS232C ports and appears to the RSX-11M operating system to be a standard input/output device. A system file is created that passes to the PDP-11/45 what the speed and characteristics of the device are. This file is run after logging-on the system and before using the VOTERM. The correct VOTERM switch settings are given in table 3 along with the RS232C connections and the contents of the initialization file.

---

[5] T.F. DeYoung, "Communication Protocols for Making the Voice Recognition Module (VRM) 'Talk' to the DG Eclipse S250 Minicomputer System," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060, report in preparation.

TABLE 2. - VRM requirements for proper operation with DG Eclipse S-250 minicomputer system running under AOS

CABLE CONNECTIONS

Pin 1 - Ground
Pin 2 - Data Out of Computer to VRM
Pin 3 - Data Into Computer from VRM
Pin 7 - Ground

DATA GENERAL SYSTEM INFORMATION

AOS Operating System
ALM Asynchronous Communications System

Device Type: 6053
Word 0: ?MEOS
Word 1: Standard
Word 2: 24 lines/page
80 characters/line

Line Initialization Word: COD3+STPO+CLK2
(where CLK2 is 2400 baud)

VRM SWITCH SETTINGS FOR DG

Switch SA: Set for 2400 baud (SA3 closed)
(N.B. All others on SA must be left open)

Switch SB: SB3: Closed
SB4: Closed
SB5: Open
(All others on SB are don't care)

Switch SC: SC5: Closed
SC6: Closed
SC7: Open
SC8: Closed
(All others on SC are don't care

Switch SD: RS232 I/O for Port 1 (Switch Label C1)
RS232 I/O for Port 2 (Switch Label C3)

Switch SE: Microphone preamplifier in (Switch Label IN)
RS232 logic levels for Port 2 (Label - PORT 2)
(N.B. When VRM is in VOTERM II, set switch to bypass preamp)

TABLE 3. - VRM requirements for proper operation with DEC PDP -11/45 minicomputer system running under RSX-11M

CABLE CONNECTIONS

Pin 1 - Ground
Pin 2 - Data Out of Computer to VRM
Pin 3 - Data Into Computer from VRM
Pin 7 - Ground

DEC PDP-11/45 SYSTEM GENERATION INFORMATION

RSX-11M Operating System
Initialization Command File:
File Name: INIT7.CMD
Contents:
SET/SPEED=TT7:2400:2400
SET/NOREMOTE=TT7
SET/SLAVE=TT7:
(N.B. The above data is for the VOTERM on Port 7 of the system; for any other port, relace the 7 with the proper port number)

VRM SWITCH SETTINGS FOR DEC PDP-11/45

Switch SA: Set for 2400 baud (SA3 closed)
(N.B. All others on SA must be left open)

Switch SB: SB3: Closed
SB4: Closed
SB5: Open
(All others on SB are don't care)

Switch SC: SC5: Closed
SC6: Closed
SC7: Open
SC8: Closed
(All others on SC are don't care)

Switch SD: RS232 I/O for Port 1 (Switch Label C1)
RS232 I/O for Port 2 (Switch Label C3)

Switch SE: Bypass preamplifier (Switch Label - OUT)
RS232 logic levels for Port 2 (Label - Port 2)
(N.B. For VRM not in VOTERM II, set switch to Preamp IN)

Once the initialization file has been activated, the VOTERM can be accessed like any other standard RS232 input/output device. That is, standard FORTRAN READ and WRITE statements are used for communicating with VOTERM. However, just as for the previous two minicomputer systems, the VOTERM commands must be predefined because the RSX-11M operating system will not be able to handle the control characters properly. The commands for this system consist of two separate words: a command word and a word for the desired function. The high order byte of the command word contains the ASCII control character DC1, the same as for the other two systems. The low order byte of the command word must have a character that the RSX-11M operating system recognizes as a valid carriage control command character. The input/output handler of the RSX-11M operating system needs it for proper operation. We chose to use the ASCII character "0" to be our carriage control character. The second data word contains the single ASCII character that corresponds to the desired VOTERM function. The responses of the VOTERM contain the ASCII control character DC2 followed by the ASCII character(s) corresponding to the commanded function. However, the RSX-11M operating system strips off the DC2, leaving only the ASCII character(s). It is therefore not necessary to predefine the VOTERM responses for this system. A more detailed description of how to use the VOTERM with the RSX-11M operating System on the DEC PDP-11/45 minicomputer is presented in another report currently in progress.[6]

The VOTALK voice synthesis system will not be discussed here because it has not been used at all on the DEC PDP-11/45. Future research plans for FY83 do include installing the VOTALK on a DEC system. The installation may be on the PDP-11/45, or it may be on the soon-to-be-installed VAX-11/780, which will be the processor for the Research Institute's Artificial Intelligence Testbed system.

The programs written to test the VOTERM on the PDP-11/45 and the one written to demonstrate what can be done with it will be discussed in sections IV and V.

---

[6]T.F. DeYoung, "Communication Protocols for Making the Voice Recognition Module (VRM) 'Talk' to the DEC PDP-11/45 Minicomputer System," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060, report in preparation.

IV. TESTING THE VIST HARDWARE

To test proper operation of the VIST hardware on the three minicomputer systems, special programs were written that would exercise all of the functions of the devices. Those for the VRM are discussed first because it was successfully tested on all three systems. The test program for the VOTALK on the HP-1000 system is covered next. As previously mentioned, the VOTALK was not fully tested on the DG Eclipse S-250 system and was not tested at all on the DEC PDP-11/45.

A. VRM Test Programs

The VRM was successfully installed and tested on three different minicomputer systems: the HP-1000 running under RTE-IVB, the DG Eclipse S-250 running under AOS, and the DEC PDP-11/45 running under RSX-11M.

The procedure for testing the VRM on all three systems was the same. First, a simple program was written that should send out the RESET command and check the VRM's response. The VRM switch settings were adjusted until this command was properly carried out. The next program reset the VRM, commanded it to go into the TRAIN mode, trained it, and then commanded it to go into the RECOGNIZE mode for the trained words. Upon successful completion of this program, the final program was written that would reset it, command it to train and recognize words, and would also command the VRM to upload (or download) the trained word data to (or from) the system's disk mass storage device. Because each system had different requirements for disk file creation and handling, they will be addressed separately.

1. HP-1000. Disk files on the HP-1000 were created using the file manager (FMGR) command CR and the FORTRAN callable routine CREAT. The files were created as Type 1 files with 128 words per record and with 100 data blocks reserved by the system for the file on a specified disk (there are 128 words per block on the HP-1000). Therefore, each record in the file contains the data for one word, that is, one VRM word pattern. This structure makes it easy to modify one word pattern without having to access the entire data file. The actual command syntax used to create the file from the FMGR was the following:

```
CR,NAMFIL:ISC:ICR:ITYPE:ISIZE
```

where: NAMFIL is the name of the file being created,
ISC is the file security code, 0 in this case
since no file security protection was desired,
ICR is the system's logical unit number of the disk,
ITYPE is the file type, a Type 1 file with 128 words
of storage per record in this case, and
ISIZE specifies that 100 blocks are reserved for the
file on the given disk

The final VRM test program for the HP-1000 had one additional feature: it could create a new disk file when the user specified a new file name. In this manner, new VRM vocabularies could be created for application programs. The HP-1000 system FORTRAN callable routine CREAT was used for this purpose. The actual syntax for the creation was the following:

```
CALL CREAT(IDCB,IERR,NAMFIL,ISIZE,ITYPE,ISC,ICR)
```

where: IDCB is a 144-word data transfer buffer,
IERR indicates whether or not the file was created,
NAMFIL is the six-ASCII-character file name,
ISIZE is a two-element array, with the 100-block file
size as the first element, and the 128-word record
length as the second element,
ITYPE is the file type, 1 in this case,
ISC is the file security code, 0 in this case since
no file security protection is desired, and
ICR is the system's logical unit number of the disk.

The files were accessed by using the HP-1000 FORTRAN callable routines OPEN and CLOSE. The files were opened, the data was transferred between the system disk and the VRM, and then the files were closed. The actual command syntax for opening the files was the following:

```
CALL OPEN (IDCB,IERR,NAMFIL)
```

where: IDCB is a 144-word data transfer buffer,
IERR indicates whether or not the file was opened, and
NAMFIL is the name of the file to be opened.

The actual command syntax used for closing the file was the following:

```
CALL CLOSE(IDCB)
```

where: IDCB is again the 144-word data transfer buffer; but it should be noted that this buffer also contains system information about the current file.

The programs written to demonstrate what can be done with the VOTERM on the HP-1000 system are discussed in section V.

2. **DG Eclipse S-250.** Disk files on the DG Eclipse S-250 were created using the AOS FORTRAN callable routine CFILW. This routine does not specify the individual record size; it specifies the type of file and the overall space to be reserved. The desired record size is specified when the file is opened. The actual syntax used to create the VRM data files was the following:

```
CALL CFILW(NAMFIL,ITYPE,ISIZE,IERR)
```

where: NAMFIL is the name of the file to be created; for the Eclipse S-250, this name consists of a six ASCII-character name with ".DAT " after the name,
ITYPE specifies the type of file, 3 for the VRM data, which signifies a contiguous file,
ISIZE specifies how many 512-byte blocks of storage to reserve, 14 for the 100-VRM word patterns, and
IERR indicates whether or not the file was created.

The files were accessed by using the DG AOS FORTRAN callable routines OPEN and CLOSE. The files were opened, the data was transferred between the system disk and the VRM, and the files were then closed. The actual command syntax for opening the files was the following:

```
OPEN LUNUM,NAMFIL,LEN=71,ERR=LABNUM
```

where: LUNUM is a logical unit number assigned to the file by the program for accessing the data on the file,
NAMFIL is the name of the file to be opened,
LEN=71 specifies that the file has a record length of 71 words, and
ERR=LABNUM specifies the program label number to jump to if an error is encountered in opening the file.

The ERR=LABNUM parameter was used to indicate that the specified file did not exist when trying to upload data to the system disk. The program would then ask if the user wanted to create the specified file and would take appropriate action, depending on the user's reply. The ERR=LABNUM parameter was used to abort the program if the download command specified a nonexistent file.

The actual command used for closing a file was the following:

```
CLOSE LUNUM
```

where: LUNUM is the logical unit number for the file on the system disk.

The program written to demonstrate what can be done with the VRM on the DG Eclipse S-250 minicomputer system is discussed in section V.

3. **DEC PDP-11/45.** Disk files on the DEC PDP-11/45 system running under the RSX-11M operating system were created using the FORTRAN callable routine OPEN. This routine was used both to create files and to open already existing files. It had a parameter called TYPE that specified whether this was a new or an old file. The actual syntax used for creating files was the following:

```
OPEN (UNIT=3,NAME=NAMFIL,TYPE='NEW',FORM='FORMATTED',)
  *ACCESS='DIRECT',RECORDSIZE=68,INITIALSIZE=14,ERR=LABNUM)
```

where: UNIT=3 is a logical unit number assigned to the file by the program for accessing the data in the file,
NAME=NAMFIL specifies the name of the file to be created; for the PDP-11/45 this name consists of a six-ASCII character name with ".DAT " on the end to tell the system this is a data file,
TYPE='NEW' indicates to the system that this is a new file to be created,
FORM='FORMATTED' specifies that the data will be formatted, i.e. a FORMAT statement will be used,
ACCESS='DIRECT' specifies that the data is to be directly accessed,
RECORDSIZE=68 specifies that each record is to contain 68 words of data,
INITIALSIZE=14 tells the system to reserve fourteen 512-word blocks for the data, and
ERR=LABNUM specifies the program label number to jump to if an error is encountered in creating the data files, in which case the program will be aborted.

The files were accessed by using the RSX-11M FORTRAN callable routines OPEN and CLOSE. The files were opened, the data was transferred between the system disk and the VOTERM, and the files were then closed. The actual syntax used for opening the VOTERM data files was the following:

```
OPEN (UNIT=3,NAME=NAMFIL,TYPE='OLD',FORM='FORMATTED',
  *                ACCESS='DIRECT',ERR=LABNUM)
```

WHERE: UNIT=3 is again the logical unit number assigned to the file by the program for accessing the data,
NAME=NAMFIL is the name of the file to be opened,
TYPE='OLD' tells the system that this is an already existing file which is to be opened,
FORM='FORMATTED' says that the data is formatted,

ACCESS='DIRECT' has the same meaning as above, and
ERR=LABNUM specifies where the program is to jump in case an error is encountered in opening the file.

The ERR=LABNUM parameter was again used to indicate that the specified file did not exist when attempting to open the file for uploading data from the VOTERM. The program would then jump to a routine that asked if the user wanted to make a new file with that name and would take appropriate action depending on the user's reply. The ERR=LABNUM parameter would abort the program if the user attempted to download data from a nonexistent file.

The actual syntax used to close VOTERM data files was:

CLOSE (UNIT=3)

where: UNIT=3 specifies the logical unit number assigned by the program of the currently opened data file.

The program written to demonstrate what can be done with the VOTERM on the DEC PDP-11/45 minicomputer system is discussed in section V.

B. VOTALK Test Programs

The VOTALK was successfully installed and tested on the HP-1000 minicomputer system. Preliminary tests were run on the DG Eclipse S-250 to determine if the VOTALK would work on this system. More extensive testing of the VOTALK on the Eclipse S-250 and initial testing of the VOTALK on the DEC PDP-11/45 may take place in FY83.

1. **HP-1000.** The procedure for testing the VOTALK on the HP-1000 system was similar to that for testing the VRM. The first test program was an extremely simple one to verify proper switch settings on the VOTALK board. This program sent the VOTALK the command to speak one of its prestored words. The VOTALK switches were adjusted until this command was properly executed. The next program tested the APPEND and PLAYBACK functions available on the VOTALK. The program would play (speak) each new sequence of specified phonemes and would append these as the next word in its user-stored vocabulary. The user could then append new words or play back any previously stored words, both those that the user stored and those prestored by the manufacturer. Several attempts were made to use the INSERT command, but so far this command has not been made to work on the HP-1000 system. The MODIFY command was tested successfully with another VOTALK test program on the HP-1000 system. That is, phonemes that were previously stored as a word were modified to improve the intelligibility of the synthesized utterance by means of the MODIFY command. The next program written to test the VOTALK will be one that will upload (or download) stored phonemes for synthesized words from (or to) the VOTALK's RAM memory to (or from) the system disk. This program is currently being written and has not yet been compiled and run on the HP-1000 system.

2. **DG Eclipse S-250.** As previously mentioned, the VOTALK has not been extensively tested on the DG system. The only program that has been run is the simplest one that commands the VOTALK to play one of its prestored words. The programs worked successfully on the DG system, but so far no additional testing has been conducted. However, plans for VIST research during FY83 include extensive testing of the VOTALK on both the DG Eclipse S-250 and on DEC minicomputer, either the PDP-11/45 or the soon-to-be-installed VAX-11/780 minicomputer system.

V. DEMONSTRATING VIST CAPABILITIES

VIST demonstration programs were written for all three of the previously mentioned minicomputer systems. However, only one of the demonstration programs includes the VOTALK voice synthesizer because the VOTALK is the most recent addition to the VIST equipment owned by CAI and it has not yet been successfully installed and tested on the other two minicomputer systems. The demonstration programs for the HP-1000 are discussed first; then the program for the DG Eclipse S-250 is described; and finally the demonstration program for the DEC PDP-11/45 is discussed.

A. HP-1000

There were two separate VIST demonstration programs written for the HP-1000. One of them demonstrates what can be done in a digital image processing environment, that is, one where the user's eyes are busy viewing an image and his/her hands are occupied moving stages and a cursor through the image. The other program demonstrates how VIST can be used to speed up the process of generating a message by using a specialized standard format, as in the DEMONS system. Although these programs demonstrate quite different tasks, the VIST requirements are similar. Because both demonstrations required the VIST equipment to perform very similar functions, a modular approach was taken with separate FORTRAN callable subroutines for these tasks. Those subroutines common to both demonstration programs will be discussed first; then the image processing program and, finally, the specialized format message generator program will be covered.

1. VOTERM Functional Subroutines. Some of the subroutines were written for specific VOTERM functions, while other routines were for program tasks that required voice input or control. The VOTERM subroutines and their purpose are given below along with a description of the necessary parameters passed by the calling program.

> VTIME(LUNR,LUTIM,IPRAM) - This routine specifies the timeout value for the VOTERM, sends the command, waits until the command has been executed, and returns to the caller.

where: LUNR is the logical unit number assigned to the VOTERM by the system,
LUTIM is the logical unit number assigned to a delay routine, and
IPRAM is the timeout value for the VOTERM in tens of milliseconds (the smallest increment allowed).

VRSET(LUVRM) - This routine resets the VOTERM and gets it ready to accept the next command.

where: LUVRM is the logical unit number of the VOTERM.

DWNLD(IFIRST,ILAST,LUVRM,NAMFIL,LUTIM) - This routine downloads the specified word pattern data from the given HP-1000 disk file to the VOTERM.

where: IFIRST is the number of the first word pattern data to be downloaded,

ILAST is the number of the last word pattern data to be downloaded,
LUVRM is the logical unit number of the VOTERM,
NAMFIL is the name of the file from which to download the data, and
LUTIM is the logical unit number of a delay routine.

VPOPN(IDCB,NAMFIL) - This routine opens the named disk file for uploading or downloading word pattern data.

where: IDCB is a 144-word data transfer buffer that also contains system information about the file, and
NAMFIL is the name of the file to be opened.

A2BIN(IASCII,IWRD) - This routine converts the two ASCII characters returned by the VOTERM into the equivalent integer for use by the calling program. The VOTERM will respond with the ASCII 'FF' when the spoken utterance is not recognized. This can cause problems unless the data from the VOTERM is read with an ASCII A2 FORMAT and then converted if the response is not 'FF'.

where: IASCII is the two ASCII characters (one HP-1000 word) returned by the VOTERM, and
IWRD is the converted integer corresponding to the ASCII characters.

STNBY(LUVRM,LUVTK,LUTIM,KNTL5) - This routine makes the calling program wait until the VOTERM recognizes the words READY and CONTINUE, then returns operation back to that program. It effectively puts the program in a standby mode awaiting these commands to continue.

where: LUVRM is the logical unit number of the VOTERM,
LUVTK is the logical unit number of the VOTALK that is used here to verify the word recognized by the VOTERM,
LUTIM is the logical unit number of the routine used to delay the program until after the VOTERM has executed the last command, and
KNTL5 is the ASCII equivalent of the VOTERM command to pass whatever follows to its auxiliary output port. For the HP-1000 VIST demonstrations this command passed data to a 32-segment LED display for user prompting and data input verification.

VHALT(LUVRM,LUVTK,LUTIM,IWORD) - This routine asks the user to verify that the program is to be halted.

where: LUVRM is again the VOTERM's logical unit number,
LUVTK is the logical unit number of the VOTALK,
LUTIM is once more the logical unit number of a delay routine, and
IWORD is the integer value of the word recognized by the VOTERM when the user is asked to verify halting the program.

2. **Image Processing Demonstration Program.** This VIST demonstration program simulates what can be done in a digital image processing environment, that is, one where the operator's eyes are busy looking at a stereoimage and his/her hands are busy moving a trackball to gather data in the stereomodel and moving a floating dot up and down for height information. The operator periodically has to interrupt the data-gathering process to input data via a keyboard. With a voice recognition device for data input and system control and a voice synthesizer for prompting and data verification, this step can be virtually eliminated, thus increasing the speed and accuracy with which data can be input and also reducing operator fatigue.

The demonstration program begins by defining the logical unit numbers for the VOTERM and VOTALK and then initializes the HP-1000's 12966A interface boards for the VOTERM and VOTALK, as previously described in section IIIA. The program then calls VTIME to set the timeout value of the VOTERM. It next defines the ASCII equivalents of the VOTERM commands and the expected VOTERM responses to these commands. The VOTERM is then commanded to reset via the VRSET subroutine, and the word pattern data is downloaded by using the DWNLD subroutine. The VOTERM vocabulary for this program is given in table 4. The VOTERM is reset again to make sure it is ready for the next sequence of commands and then it is put in the recognize mode. While the VOTERM is in this mode, there are two separate ranges of words that are compared for valid utterances. One set of words are those numbered 0 through 5 in table 4. These are the so-called 'COMMON' vocabulary, that is, they are always to be checked for valid utterances. The other allowable word range contains the words that are specific to a given situation. This range is changed throughout the program by telling the VOTERM the new range of valid words to compare with the spoken word.

The first of the 'COMMON' words, STANDBY, is used to call subroutine STNBY to put the program in a wait state until the two 'COMMON' words READY and CONTINUE are given in sequence. The program then continues normal execution. The command to HALT causes the program to call the VHALT subroutine, which prompts the operator to verify halting the program. If the operator says HALT, the program stops, while the command CONTINUE will cause the program to continue normal execution. The 'COMMON' word RETURN makes the program branch to the previous menu of available commands and prompts the operator for input.

The operator is prompted to say one of the valid words by a menu on the CRT screen by displaying the valid words in a 32-character format on the the 32-segment display, and by saying the words with the VOTALK voice synthesizer. One of the 'COMMON' words is always LIST, which outputs all three types of prompts again. This is a very useful word if the operator forgets where he/she is in the program. The 32-segment display simulates what could be displayed in one of the eyepieces of a stereoviewer using the superposition capabilities of a modern image-processing system like the CAPIR.

Depending on the operator's spoken command, the demonstration program will execute the corresponding part of the code. The CURSOR and SCROLL IMAGE commands branch to routines that enable the operator to roam through the image. The operator can either use the trackball to drive the motion or input the desired motion by voice command. A voice command enables the user to alternate between trackball-driven motion and voice-commanded motion.

TABLE 4. - VOTERM recognition vocabulary for image processing VIST demonstration program on HP-1000 minicomputer

NUMBER	WORD	NUMBER	WORD
0	STANDBY	21	ONE
1	HALT	22	TWO
2	RETURN	23	THREE
3	LIST	24	FOUR
4	READY	25	FIVE
5	CONTINUE	26	SIX
6	SCROLL IMAGE	27	SEVEN
7	CURSOR	28	EIGHT
8	DIGITIZE	29	NINE
9	UP	30	BACKSPACE
10	DOWN	31	CANCEL
11	RIGHT	32	MARK
12	LEFT	33	CLOSE BOUNDARY
13	DIAGONALLY	34	END FEATURE
14	ZOOM-IN	35	ENTER DATA
15	MOVE OUT	36	IMAGE
16	HOME	37	OVERLAYS
17	POINT	38	DISPLAY ALL
18	LINE	39	VOICE COMMAND
19	AREAL	40	MANUALLY
20	ZERO	41	STOP

The DIGITIZE command branches to a routine that enables the operator to input new data by voice or to change the display according to the spoken display option. This particular part of the image-processing demonstration program most closely simulates how VIST can be used in a system like CAPIR. The ENTER DATA command branches to a routine used to simulate the data digitization mode on the CAPIR system. During the data entry mode the numbers 0 thru 9 and the edit commands BACKSPACE and CANCEL are valid utterances. The last two commands are used to correct any erroneous data, whether due to operator error or to a misrecognition of the VOTERM. The operator is first prompted to enter the Facility Access Code (FAC) number for the data to be digitized. Once the FAC number is given, the user is prompted to input the type of feature being digitized, either POINT, LINE, or AREAL. The MARK command is then used to enter the current position of the cursor for the given FAC number. The MARK command terminates data entry for a point feature, while the END FEATURE and CLOSE BOUNDARY commands are used to terminate digitizing line and areal features, respectively.

The VIST demonstration program described above shows that a voice recognizer coupled with a voice response unit can be used on an image-processing system to improve the speed and accuracy with which data can be input and at the same time reduce operator fatigue caused by eyestrain.

3. **DEMONS Demonstration Program.** This VIST demonstration program simulates what can be accomplished in an environment like that on the DEMONS system. The program simulates two different users of the DEMONS system: an image supervisor who calls up frames of imagery and designates certain areas for an analyst to examine and an image interpreter who displays these areas and generates standardized format messages about what is in these areas.

The DEMONS demonstration program currently does not use the VOTALK voice synthesizer, but plans for FY83 include adding this capability to the demonstration.

TABLE 5. - VOTERM recognition vocabulary for DEMONS VIST demonstration program on HP-1000 minicomputer

NUMBER	WORD	NUMBER	WORD
0	STANDBY	40	MATRIX
1	HALT	41	ARMOR
2	RETURN	42	ARTILLERY
3	LIST	43	ARMORED PERSONNEL CARRIER
4	READY	44	UNKNOWN
5	CONTINUE	45	HEAVY
6	FRAME	46	LIGHT
7	ROAM	47	T-72
8	TRANSFER	48	T-55
9	STOP	49	TANK
10	ZERO	50	ZSU 23-4
11	ONE	51	ANTI-AIRCRAFT ARITLLERY
12	TWO	52	FIELD ARTILLERY
13	THREE	53	SELF-PROPELLED
14	FOUR	54	TOWED
15	FIVE	55	LAUNCHER
16	SIX	56	MISSILE
17	SEVEN	57	ROCKET
18	EIGHT	58	IN FIRING POSITION
19	NINE	59	MOVING
20	BACKSPACE	60	NORTH
21	CANCEL	61	SOUTH
22	UP	62	EAST
23	DOWN	63	WEST
24	RIGHT	64	NORTHEAST
25	LEFT	65	SOUTHEAST
26	DIAGONALLY	66	NORTHWEST
27	ZOOM	67	SOUTHWEST
28	MOVE OUT	68	AT TIME
29	HOME	69	NO CHANGE
30	CENTER	70	NOT SIGHTED
31	MARK TARGET	71	ENTER TEXT
32	ALPHA	72	CORRECTION
33	BRAVO	73	NUMBER
34	DISPLAY	74	DESCRIPTION
35	HOT	75	ACTIVITY
36	UNCLASSIFIE	76	LOCATION
37	CONFIDENTIAL	77	REMARKS
38	SECRET	78	SUPERVISOR
39	TOP SECRET		

1. **VRM Functional Subroutines.** There were two VRM functional subroutines written for the DG Eclipse S-250: the first one initialized the VRM, and the other one put it in the recognition mode for the designated words and returned the number of the word recognized. The routines are given below with a description of their purpose and the necessary Parameters passed by the calling program.

VOINIT(NAMFIL,NFIRST,NLAST) - This subroutine initializes the VRM by resetting it, setting the rejection level for nonrecognition, and downloading the word pattern data to the VRM from the named disk file.

where: NAMFIL is the name of the disk file from which to download the word pattern data,
NFIRST is the number of the first word pattern data to be downloaded and,
NLAST is the number of the last word pattern data to be downloaded.

VOISIN(NFIRST,NLAST,NWRD) - This subroutine reset the VRM, put it in the recognize mode for the designated range of words, took the ASCII response from the VRM, converted it to an integer, and returned this number to the calling program.

where: NFIRST is the number of the first word to check for a valid utterance,
NLAST is the number of the last word to compare for a valid utterance, and
NWRD is the number of the word recognized by the VRM.

2. **Vertical Obstruction (VO) Demonstration Program.** This VIST demonstration program shows how voice data input can be used in an environment like the VO task on the CAPIR system, where the operator's eyes and hands are busy. Without the voice input capability, the operator periodically has to turn around to input information by means of the keyboard. This process can cause operator fatigue due to eyestrain and can slow the rate of data input into the system. If the system has the capability for voice data input, this step can be virtually eliminated.

As previously mentioned, the VIST VO demonstration was part of a program to show how the CAPIR system could be used to digitize VO data. The VIST program was a subroutine that was called when the operator chose the voice input option from a menu on the CRT screen. The VIST routine started by calling the VOINIT subroutine to intialize the VRM and download the proper word pattern data to the VRM. The vocabulary for the VIST VO program is given in table 6. The program proceeded to prompt the operator to digitize the VO by using the normal digitization routine and by using the footpedal for storing the data for the bottom and the top of the object. The program then asked the operator to say what type of VO was being digitized. The operator then said one of the following: RADIO TOWER, POWER LINE, FLAGPOLE, BUILDING, or TREE. The program then displayed the location and height of the object and what type of VO it was on the CRT screen. The operator was then prompted to say either NEXT to continue digitizing other VO objects or RETURN to go back to the main menu of options. It should be noted that this VIST demonstration is not nearly so complex as those for the HP-1000, but it nevertheless shows how VIST can be used effectively on the CAPIR system to reduce operator fatigue due to eyestrain and at the same time improve the data input rate.

TABLE 6. - VRM recognition vocabulary for vertical obstruction demonstration program on DG Eclipse S-250 minicomputer

NUMBER	WORD
0	NEXT
1	RETURN
2	RADIO TOWER
3	POWER LINE
4	FLAGPOLE
5	BUILDING
6	TREE

C. DEC PDP-11/45

There was one VIST demonstration program written for the DEC PDP-11/45 system. This program showed how a voice recognition device could be used to manipulate graphics displayed on a color monitor. The VIST equipment used for this demonstration was the VRM-102 board in the VOTERM chassis.

Although there was only one VIST demonstration program written for the DEC PDP-11/45 system, a modular approach was used to create the routine that initializes the VOTERM and downloads the word pattern data to it. This routine may therefore be used by any future VIST programs. This functional subroutine will be discussed first and then the graphics manipulation by voice program will be covered.

1. VOTERM Functional Subroutine. The VOTERM functional subroutine written for the DEC PDP-11/45 system performed the same function as the initialization subroutine for the DG Eclipse S-250. The routine is given below, along with the parameters passed by the calling program.

> VOINIT(NFIRST,NLAST,NAMFIL) - This subroutine initialized the VOTERM by resetting it, setting the rejection level threshold for nonrecognition, and downloading the specified word patterns from the designated disk file.

where: NFIRST is the number of the first word pattern data to be downloaded,
NLAST is the number of the last word pattern data to be downloaded, and
NAMFIL is the name of the disk file from which to download the data.

2. **Graphics Manipulation Program.** This VIST demonstration program shows how graphics displayed on a color monitor can be manipulated by using voice control. The program uses the VOTERM voice recognition device to access different images stored on a system disk, display them on a color monitor, and display various overlays on these images.

The first thing that the program does is to call subroutine VOINIT to initialize the VOTERM and download the proper word pattern data needed for this demonstration. The vocabulary for this program is shown in table 7. The program will then display a menu of available commands. If the operator says RETURN or QUIT, the program is halted. The IMAGE command will cause the program to branch to the image portion of the code and a new menu of available commands is displayed on the CRT screen. The valid commands here are SOUTH, NORTH, and LOWER. These correspond to a digitized representation of Cache, Oklahoma, from the South, North, and a view closer to the ground from the North. After the image is displayed, the OVERLAYS command puts the operator in the overlay mode, where individual overlays can be turned on or off. That is, if the overlay is on, saying the name again will turn it off. The available commands for this portion of the program are words 7 thru 12 in table 7. The RETURN and QUIT commands used here will cause the program to branch back to the first menu of available commands.

Although this VIST demonstration program is not nearly so complex as those written for the HP-1000, it nevertheless shows how a voice recognition device like the VOTERM can be used effectively to manipulate graphics displayed on a color monitor and how different images can be recalled from a disk and displayed by voice command.

TABLE 7. - VOTERM recognition vocabulary for graphics manipulation program on DEC PDP-11/45 minicomputer

NUMBER	WORD
0	RETURN
1	QUIT
2	IMAGE
3	OVERLAYS
4	SOUTH
5	NORTH
6	LOWER
7	GRID
8	CASTLE
9	RANGE
10	HEAVY DUTY ROADS
11	LIGHT DUTY ROADS
12	HYDROGRAPHY

VI. RESULTS AND CONCLUSIONS

The results to date have been very encouraging. The VIST demonstration programs have shown that voice data input and control can be used effectively when the operator's eyes and hands are busy, as in the CAPIR and DEMONS systems. As a result of these demonstrations, it was decided to install a voice recognition capability on the DEMONS and CAPIR systems. A contract has been let to install a VOTERM voice recognizer on the DEMONS system. ETL personnel will act as consultants to the contractor to insure a successful transfer of VIST to this project. An unsolicited proposal was received for installing a voice recognition and synthesis capability on the CAPIR system, and it was recommended to fund this project. Work on this effort should begin during the first quarter of FY83. ETL personnel will again act as consultants to insure a successful transfer of VIST to this project.

Although the results have been good, a great deal of work remains in order to make these systems more user-friendly. One way that this can be done is by using a natural language interface between the operator and the system. As a first step towards this natural language understanding system, it is recommended that a limited connected-speech recognition system be purchased. The current voice recognizer requires that each word be spoken with a pause before the next word so that it can process the spoken utterance. This type of recognizer is an isolated-word recognizer. The limited connected-speech recognizer permits the operator to speak a sentence of four or five words before the device processes the utterances and causes the system to take appropriate action. Clearly, this latter type of device is more desirable because it more closely resembles a natural language understanding system.

Another way to improve the user interface would be to have a more natural sounding speech synthesizer. The current device uses phonemes to create the speech, with the result that the sound is somewhat mechanical and is a little difficult, at first, to understand. The main advantage of this type of device is that it has an unlimited vocabulary. However, with the amount of storage available on a single RAM chip going up all the time, the stored speech synthesis devices become more attractive, especially in a situation where the vocabulary can be limited to a designated set of words. This is precisely the case in the DEMONS and CAPIR systems. It is therefore recommended that a stored-speech development system and output device be purchased and incorporated

into these systems as soon as funds are available. Prior to their incorporation in these systems, it is further recommended that they be installed and tested on the HP-1000 minicomputer in the CAI to determine how this type of device might best be used by the DEMONS and CAPIR systems.

Another potential use for VIST would be in the Terrain Analysis Demonstrator System (TADS) that ETL has proposed as part of the Army's AI/Robotics program for FY83 and FY84. The VIST hardware would be used in the TADS to command the display system to output desired terrain and topographic information on color monitors to the right and left of an operator who is teleoperating a vehicle via a UHF radio link. VIST has been included in this fashion in the proposed research plan for the TADS.

Future plans for the VIST research at ETL include purchasing the above-mentioned hardware and installing the current VIST hardware on the CAI's new AI Testbed system, which has a DEC VAX-11/780 superminicomputer as its central processor. Other systems under development at ETL will be investigated to determine if VIST can be effectively used to improve their performance.

VII. REFERENCES

DeYoung, T.F., "Preliminary Speech Recognition/Synthesis Experiments," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA, September 1980, unpublished report.

DeYoung, T.F., "Communication Protocols for Making the Voice Recognition Module (VRM) 'Talk' to the HP-1000 Minicomputer System," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA, 1982, report in preparation.

DeYoung, T.F., "Communication Protocols for Making the Voice Recognition Module (VRM) 'Talk' to the DG Eclipse S-250 Minicomputer System," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA, report in preparation.

DeYoung, T.F., "Communication Protocols for Making the Voice Recognition Module (VRM) 'Talk' to the DEC PDP 11/45 Minicomputer System," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA, report in preparation.

Lukes, G.E., "Computer-Assisted Photo-Interpretation Research (CAPIR): A Prospectus," U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA, November 1979, unpublished report.

Wildes, A.R., DVBOO, A General Purpose, Multiple Device Driver for the HP 12966A Buffered Asynchronous Data Communications Interface, NASA, Goddard Space Flight Center, Greenbelt, MD, March 1977.

FILMED

8