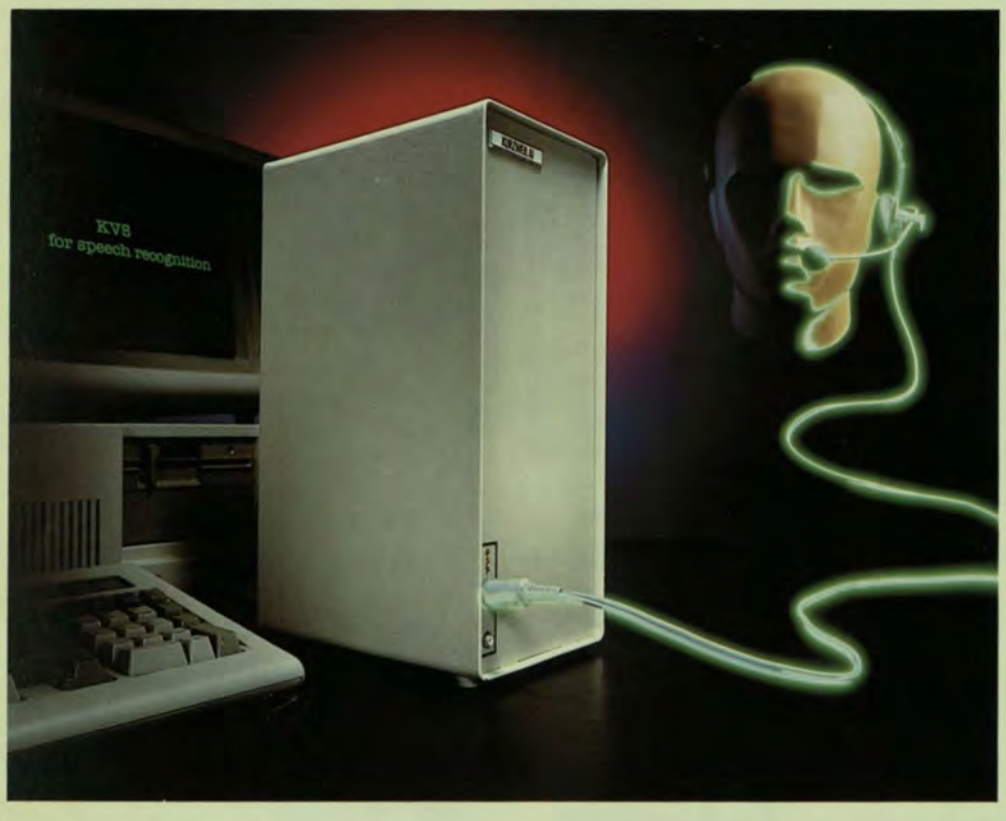
KURZWEIL AI

KURZWEIL AI

Talk Instead of Type



VOICE INPUT: MAKING COMPUTERS LEARN OUR LANGUAGE

We are on the threshold of a new era in computer technology, that will mean profound changes in the way information is stored, processed, and communicated. It is an era when people will control computers directly by voice.

Voice is probably the most effective means of communications for getting work done quickly and efficiently. Now, Kurzweil Applied Intelligence's developments in voice input allow people—not just proficient typists and computer technicians—to operate computers and generate information simply by talking.

The Complexities and Challenges of Voice Recognition

Talking, in person or by phone, is an accepted way of getting and giving information. Sometimes, we take it for granted because it's so natural. But the ability to understand the sounds and inflections of speech is a highly complex human skill, and designing machines that mimic this trait has been a challenging process.

Demonstrating feasibility in the laboratory is difficult enough. The challenge is even greater when the machine must meet the additional needs of commercial markets including: large vocabulary size, high accuracy, simple human interface, low cost, and compact physical size.

Accepting the Challenge

In 1982, Raymond Kurzweil, an established expert in pattern recognition and speech processing, assembled a team of specialists from the fields of acoustic science, linguistics, signal processing, software development and hardware engineering. The goal: to design an accurate, commercially-feasible way to create text by voice.

The strategy was set at the very beginning. Rather than pursue a single, narrowly-focused technique, the multi-disciplinary team combined the best features of several approaches—acoustic-phonetic, statistical, and linguistic. This multiple expert procedure would incorporate sophisticated artificial intelligence techniques, involving millions of separate calculations, that must be performed in a fraction of a second. Each word would appear on the screen as it is spoken.

Voice Enhances Productivity

PRACTICAL APPLICATIONS

Kurzweil products have been used in a variety of commercial applications, from small operations to Fortune 500 companies. Their advanced capabilities have opened new applications requiring large vocabulary voice recognition.

Text Creation and Editing

A fundamental office activity, creating written text and documents is usually done by handwriting, dictating, or typing. All of these methods prove inefficient when compared with voice recognition.

Many people write documents in longhand, a process ten times slower than voice dictation, and a real productivity bottleneck. Or, they may dictate into a tape recorder, or to a secretary. However, there is a scarcity of secretaries skilled in taking rapid dictation. The person dictating text often has difficulty organizing and correcting copy without seeing it, resulting in reviews and rewrites. Furthermore, many professionals lack access to dictation services.

A person can type a document directly into a typewriter or word processor. But many people do not have (or want to learn) typing skills. Even for those who can type, voice is still substantially faster, and more natural.

Voice input allows generation of printed text, without the transcription and editing steps required with handwriting or dictation. Benefits include quicker turn-around, with less demands on support staff.

Database Inquiry and Analysis

Through simple spoken commands, professionals can get the information they need from the computer when they need it. Access to information is no longer limited to skilled typists, and programmers. Faster access to data can mean better decisions, and reduced demand for support staff translates into time and money savings.

Division managers at Fortune 500 companies use voice to obtain product and financial information, when needed.

Government agencies use voice to increase productivity by reducing overhead. Senior personnel can access information from large databases quickly, without relying on support staff.

Data Entry and Translation

Data entry is a basic function that cuts across a wide range of industries and markets. Computer output can be only as accurate and timely as the information entered. In some organizations, the volume of data that must be entered is staggering—sometimes, millions of words or codes daily.

At insurance and large mail order companies, data entry operators rapidly sort through a stack of forms, skimming over each one to find a certain field or box, and simply speaking the information into the computer. Voice input virtually eliminates typographical errors, since translation of each spoken word is pre-defined and automatic. Complex look-up codes are eliminated. Furthermore, voice reduces the amount of training, and reduces operator fatigue.

Direct voice input allows pathologists to enter data directly into a computer, even while looking through a microscope, or performing an autopsy.

Manufacturing quality control inspectors enter data quickly and accurately using voice, which frees their hands to hold parts, and their eyes to perform visual inspections.

KURZWEIL AI

Command and Control

In many scientific, technical, engineering, and industrial environments, voice is being used to control workstations, machines and robots.

Voice complements keyboard and mouse input in technical workstation applications:

- Computer-aided Design (CAD)
- Computer-aided Engineering (CAE)
- Scientific and Industrial Data Acquisition
- Electronic Printing and Publishing

Voice input is easier, faster, and more accurate than keyboard or mouse entry, especially for menu selection. Productivity gains from 20–30% have been measured, derived primarily from better concentration on the task.

Voice is being used to control machines and robots. In these applications, the operator typically cannot access a keyboard. The user may already be operating a machine on the manufacturing floor, or may be handicapped and using voice to control wheelchair movement.

For the physically handicapped, voice recognition offers real hope for bettering personal communications, education, and gainful employment. Voice input drives word processors, spreadsheets, database management packages, programming languages, and telecommunications systems, helping the handicapped realize their full potential value to our society through productive employment.

MARKETS

The application of voice recognition has impact across a number of markets and industries. Kurzweil is working with Value-added Resellers (VARs) and end-users to develop solutions to real-world problems by using large vocabulary voice recognition technology.

Office

- General purpose text creation
- Legal documentation
- Customer service
- Insurance forms data entry
- Financial services
- Brokerage data query
- Desk-top publishing

Medical

- Radiology reporting
- Pathology reporting
- Patient records
- Nursing records
- Handicapped command and control

Government

- Aerospace operations
- Command and control of vehicles
- Office automation
- Natural language database inquiry
- Intelligence data entry

Telecommunications

- Command and control of equipment
- Remote data access by telephone
- Office automation
- Telemarketing

Engineering

- CAE/CAD/CAM
- Technical publication
- Software development

Manufacturing

- Data entry of inventory, parts, quality inspection
- Data inquiry of MRP systems
- Communication with AI expert systems
- Command and control of machines, and robots

APPLIED TECHNOLOGY

Kurzweil's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering production costs and prices over time.

In this information-dependent age, the need to create text, enter data, and analyze information has become an integral part of the way we work. The keyboard, however, has been a major obstacle to those who lack typing and computer skills. Ironically, those who most need the information often can't get at it.

The advent of voice recognition technology expedites the flow of information in computer processing. It provides a new degree of efficient control over everything computers do, from record-keeping and word processing to controlling manufacturing processes. Now, anyone can enter and retrieve information easily. With the removal of the keyboard barrier, the market potential for Kurzweil AI technology seems virtually unlimited.

KURZWEIL AI

Voice Made Practical

KURZWEIL AI PRODUCTS

Kurzweil Voiceworks™ KVW

The Kurzweil Voiceworks (KVW) is the first practical, commercially available voice recognizer capable of tackling the challenging demands of general purpose text creation. Until now, voice recognition systems have been limited by vocabularies that are too small to support applications such as ordinary dictation.

Kurzweil AI's speech scientists and engineers have developed and applied advanced pattern recognition and artificial intelligence techniques to produce the KVW, a voice recognition system with a working vocabulary of between 5,000 and 10,000 spoken words or phrases. The KVW features real-time response with the degree of accuracy and reliability required for commercial applications.

The primary application of the KVW is to automate the creation of written text, which is a fundamental activity in the office. After users train the KVW to understand their speech patterns, they can simply dictate text and immediately see the words appear on the screen as they are spoken.

The KVW's large vocabulary and flexibility make voice input and control possible for computer-based applications in business, medicine, legal, publishing, and government.

Kurzweil Voicesystem™ KVS

The Kurzweil Voicesystem (KVS), Kurzweil AI's first voice recognition product with vocabularies of up to 1,000 words, has been in production since mid-1985. The KVS was the first commercially available voice recognizer with sufficient vocabulary to allow direct voice control and voice input for many industrial and business systems, at a cost that opened up a host of new application possibilities.

Until the introduction of the KVS, voice recognition systems had been hampered by limited accuracy, insufficient vocabulary size, excessive sensitivity to ambient sound, and high initial cost. With hundreds of units installed during its first year of shipment, the KVS has been incorporated into diverse data management, command-and-control, and limited text creation applications.

Kurzweil Voiceterminal™ KVT

The Kurzweil Voiceterminal™ KVT builds upon the KVS technology and functions as an intelligent voice-driven terminal that enables people to use spoken words and phrases to control, enter data into, and retrieve data from mainframes and minicomputers. The KVT incorporates this technology into a terminal that communicates with IBM, DEC, HP or other host systems, without modification of the host application.

In addition to running mainframe applications with voice input, the KVT provides IBM PC/XT compatibility with support for most popular IBM PC applications. Given this capability, the KVT can also be used in standalone mode, without any connection to a host system.

Kurzweil VoiceRAD™ KVR

The Kurzweil VoiceRAD system is designed for dictating radiology reports directly to a computer. Using the KVR, a doctor can read films, dictate, print and sign reports in one efficient session. Time-saving standard phrases trigger the typing of entire paragraphs. Once the report is ready, the radiologist simply tells the system "print this". The system immediately produces an accurate printed report for updating the patient's records and distribution to the health care team.

COMPANY HEADQUARTERS

Kurzweil Applied Intelligence Incorporated
411 Waverley Oaks Road
Waltham, MA 02154-8465
617-893-5151 Telex: 9103809408

Sales Offices

East

Boston: 411 Waverley Oaks Road
Waltham, MA 02154
(617) 893-5151

New York: 120 Wood Avenue South
Iselin, NJ 08830
(201) 321-0555

Wash., D.C.: 6411 Ivy Lane
Greenbelt, MD 20770
(301) 982-0666

Central

Chicago: Suite 400
8410 W. Bryn Mawr Avenue
Chicago, IL 60631
(312) 380-1215

Dallas: Infomart
Suite 5001
1950 Stemmons Freeway
Dallas, TX 75207
(214) 746-5023

West

N. Calif.: Suite 200
1301 Shoreway Road
Belmont, CA 94002
(415) 591-1708

S. Calif.: Suite 1100
2049 Century Park East
Los Angeles, CA 90067
(213) 203-9909

International

Contact company headquarters for international distributor locations.

Printed in USA K12.8620M2.0

KURZWEIL VOICEWRITER™

KURZWEIL AI

VOICE RECOGNITION COMES OF AGE

Voice is our most effective means of communication. Compared to keyboard input for data entry, database inquiry, and text creation, voice input is much faster while requiring less training.

The Kurzweil Voicewriter (KVW) is the first practical, commercially available speech recognizer capable of tackling the challenging demands of general purpose text creation. Until now, voice recognition systems have been limited by vocabularies that are too small to support applications such as ordinary dictation.

ADVANCED TECHNOLOGY

Kurzweil AI's speech scientists and engineers have developed and applied advanced pattern recognition and artificial intelligence techniques to produce the KVW, a speech recognition system with a working vocabulary of between 5,000 and 10,000 spoken words or phrases. The KVW features real-time response with the degree of accuracy and reliability required for commercial applications.

The technology incorporated into the KVW is complex, but its basic function can be described quite simply. The KVW recognizes spoken words or phrases and translates them into strings of ASCII characters. The resulting character strings can be interpreted as keyboard input for virtually any computer-based application.

VOICE-ACTIVATED WORD PROCESSOR

The primary application of the KVW is to automate the creation of written text, which is a fundamental activity in the office. When coupled with a word processor, the KVW provides general voice-to-text creation. After users train the KVW to understand their speech patterns, they can simply dictate text and immediately see the words appear on the display screen as they are spoken.

Kurzweil is working closely with major vendors of word processing and office automation systems to provide tightly integrated system solutions. This integration will allow users not only to compose text by speaking, but also to display, edit, process, store, and print documents by simply speaking the appropriate commands.

The system starts with a vocabulary of at least 5,000 frequently used words in the English language. The KVW adapts to each individual's own vocabulary. The system subsequently adds words that are not part of its initial vocabulary and eventually deletes those words that are never used. User-specific vocabularies containing the most frequently used words for a given application area, such as medicine or law, will also be offered, so that users will enjoy the immediate benefits of having a large initial working vocabulary included in the system.

INTEGRATED APPLICATIONS

The KVW is not limited to word processing applications. For example, the KVW can be combined with natural-language understanding to enable professionals to make database inquiries verbally in natural language instead of through a keyboard.

The KVW's large vocabulary and flexibility make voice input and control possible for computer-based applications in business, medicine, government, defense, and education.

The KVW is suitable for commercial and industrial applications including:

- ☐ memo and report dictation
- ☐ medical record transcription
- ☐ legal document preparation
- ☐ printing and publishing
- ☐ natural language database inquiry
- ☐ communication for the handicapped

The KVW provides an opportunity for innovative companies to integrate voice with their applications and benefit from the improved productivity and throughput.

To facilitate development of volume applications, Kurzweil AI is prepared to work closely with end-user customers, value-added resellers, and value-added distributors.

KURZWEIL VOICEPARTNER PROGRAM

The Kurzweil Voicepartner Program, established initially for the Kurzweil Voicesystem™ (KVS), provides a broad range of applications from leading software suppliers. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications. Building upon the base of Voicepartners who already support the KVS, Kurzweil AI will continue to expand its Voicepartner program to support new applications, based upon the KVW's additional features.

Multiple Kurzweil Voicewriter user stations can be connected to a single voice server.

KVW BENEFITS

- ☐ **Ease of operation**—as natural as talking
- ☐ **Accuracy**—no typographical errors
- ☐ **Time economy**—reports can be produced immediately
- ☐ **Productivity**—can be used by non-typists
- ☐ **Speed**—rapid entry of text and commands
- ☐ **Hands-free**—work and dictate simultaneously

KVW PRODUCT FEATURES

- ☐ **Large vocabularies**—between 5,000 and 10,000 phrases/words can be recognized at any given time without switching vocabularies.
- ☐ **User-definable vocabulary**—new words can be easily and quickly added to the vocabulary.
- ☐ **Real-time response**—each word is displayed on the monitor as it is spoken. Less than a 1/4 second pause is required between words.
- ☐ **Simple human interface**—does not require prior computer experience.
- ☐ **English-based Language Model:**
  –applies artificial intelligence concepts to maximize accuracy
  –properly recognizes homonyms based upon sentence context (e.g. to, two, too, 2)
- ☐ **Multiple vocabularies**—users may access multiple vocabularies of over 5,000 words each.
- ☐ **Multiple users**—several user stations can be connected to a single KVW voice server, providing the lowest cost per user. The voice server supports one user at a time.
- ☐ **Spoken commands**—can be recognized during dictation without using the keyboard. A single speech command can replace a complex string of keyboard strokes.
- ☐ **Highly accurate and reliable**—the KVW is trained by each speaker, providing the accuracy needed for many business and industrial applications.
- ☐ **Self-contained, compact-sized unit**—fits neatly on a desk.
- ☐ **Software development tools**—for the system integrator to enhance integration of KVW with computer-based applications, and support non-IBM PC/AT compatible host systems.

SPECIFICATIONS

Voice Server

Host Configuration: IBM PC/AT

Power:
110 VAC-50/60 Hz at 7.0 Amps
220 VAC-50/60 Hz at 3.5 Amps

Dimensions: 16″h, 8″w, 13″d

Weight: Less than 35 lbs.

Voice Station

Host Configuration: IBM PC/XT/AT or compatible

Dimensions: 14″h, 6.5″w, 8″d

Weight: 18 lbs.

OPERATING CONDITIONS

Temperature: 10-40 C (50-104 F)

Relative Humidity: 10-90%

PRODUCT FAMILY

The Kurzweil Voicesystem (KVS), Kurzweil AI's first voice recognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry, applications.

The Kurzweil Voicewriter (KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of handling virtually any printed text and converting it to synthesized speech. In 1984, he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly re-create the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development, and engineering. The company's technological leadership is backed by a commitment to full technical support and service, worldwide. Kurzweil engineering and marketing staff work closely with customers to adapt Kurzweil AI products to specific applications.

KURZWEIL AI

For further information on the KVW and other Kurzweil AI products, contact your nearest sales office.

COMPANY HEADQUARTERS

411 Waverley Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverley Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Suite 161, 2915 L.B.J. Freeway, Dallas, TX 75234, (214) 247-3622

West

N. Calif.: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif.: Suite 1000, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

Suite 800, 685 Town Center Drive, Costa Mesa, CA 92626, (714) 641-3202

International

Contact company headquarters for international distributor locations.

KVS PORTING MANUAL

OVERVIEW

The Kurzweil Voicesystem ™ (KVS) Porting Manual contains software source programs (disks and hardcopy listings) and documentation to allow software engineers with experience in writing device drivers to port the KVS host software to other operating systems and/or host hardware. Familiarity with the KVS is assumed. The manual describes the design, testing, and operation of the KVS host-resident software by using Kurzweil's current implementation on the IBM PC/XT/AT running PC-DOS as an example. In one chapter, guidelines for porting to Sun Microsystem's UNIX are described to further illustrate the porting concepts.The software that runs within the KVS itself is not modified during the porting process.

The KVS Porting Software consists of three separate parts, all of which run on the host processor:

- KVS Device Driver.
- Run-time Library.
- Host Utilities.

Each of these modules must be ported to provide KVS support on other computer systems.

KVS DEVICE DRIVER

The KVS Device Driver works with the host operating system to provide support for the KVS.

The KVS Device Driver interacts with the following devices, which are found in most computer systems:

- Display
- Disk
- Serial communications
- Keyboard buffer.

The functions provided by the KVS Device Driver include:

- support of the communications hardware and protocol
- feeding KVS output to the application program running on the host
- mass storage of phrasebooks via operating system service calls at user request
- transparent keyboard service. The keyboard handler is modified to send all keystrokes to the KVS for filtering and echoing back to the host
- window management services, providing a user and application program access to the KVS pop-up window
- allowing the KVS to access operating system services
- maintaining the host's type-ahead buffer
- presenting an interface to a voice application program providing specific entry points for reading and writing messages to/from the KVS, checking status and controlling I/O.

The KVS Device Driver, available from Kurzweil AI for the IBM PC/XT running PC-DOS, is written in assembly language. Ports to other computer systems can be based upon this existing code.

RUN-TIME LIBRARY

The Run-time Library is a collection of routines callable from a voice application program which handles interactions between an application, the KVS, and the language model. These routines allow applications programmers to tailor their applications to the KVS. The Run-time Library can be used to:

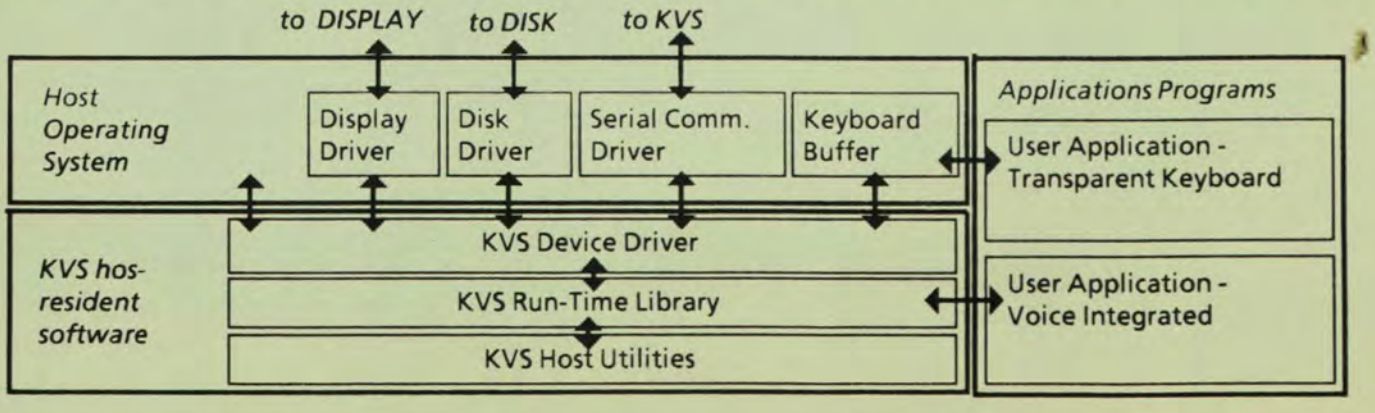
- **Interface the KVS to other computers or devices connected via the host computer.**

  The Run-time Library includes the tools to develop the required custom voice application programs.

- **Process words prior to display.**

  In transparent mode, spoken words which are recognized by the KVS are passed to the application program and displayed simultaneously on the screen. Using the Run-time Library, these words can be processed by an application program so that contextual information can be added by the application prior to the display of the characters.

Relation between Host Operating System, Application Programs, and KVS host-resident software



- **Customize the human interface.**

  The windows displayed can be modified and special help messages can be added to suit the user's need.

The Run-time Library is not hardware specific and contains few operating system specific sections. Written in Lattice C ™, the Run-time Library can easily be ported to many other systems.

For programmers who already have a host machine with KVS support, the Run-time Library is available separately, without the driver and host utilities sources that are packaged with the Porting Manual.

HOST UTILITIES

Host Utilities run provide a variety of user services including:

- validating user data
- providing help screens
- verifying sufficient disk space for enrollment and phrasebook creation
- downloading phrasebooks and programs
- sending angle bracket commands to the KVS for the user

Like the Run-time Library, the Host Utilities are not hardware specific and contain few operating system specific sections. They are written in Lattice C to facilitate porting to other systems.

SYSTEM CONFIGURATION

Host Requirements:

- Display-to-host connection.
- RS-232 port for KVS connection. Support of baud rates higher than 9600 baud is recommended to minimize vocabulary loading time.
- 1 Mbyte incremental hard disk storage for KVS Porting Software.
- 20 Kbyte incremental RAM storage for KVS Device Driver (may vary with each port).

Software Requirements:

- Lattice C (not supplied with Porting Manual).
- Microsoft Macro Assembler (for 8086-based ports only).

SUPPPORT/WARRANTY

Training, telephone support, and on-site support are available. 90 day warranty. Post warranty updates will be available for an additional charge.

ORDERING INFORMATION

KVS-S003	KVS Porting Manual (sources and documentation), Training, One Day Telephone Support
KVS-S004	KVS Porting Manual (sources and documentation), Training, One Week On-site Support
KVS-S002	KVS Run-time Library object code (does not include KVS Device Driver, Host Utilities, and Run-time Library sources)

For further information on the KVS Porting Manual and other Kurzweil AI products contact your nearest sales office.

KURZWEIL AI

COMPANY HEADQUARTERS
411 Waverley Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverley Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Ave., Chicago, IL 60631, (312) 380-1215

Dallas: Suite 161, 2915 L.B.J. Freeway, Dallas, TX 75234, (214) 247-3622

West

N. Calif.: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif.: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

Suite 800, 685 Town Center Drive, Costa Mesa, CA 92626, (714) 641-3202

International

Contact company headquarters for international distributor locations.

Data Entry:
Voice Input Makes Data Entry More Productive

Computer output is only as accurate and timely as the information entered. The volume of data that businesses must enter daily to feed and update is staggering -- in some companies, millions of words or codes just to maintain their business.

The job is tedious, repetitive and often redundant. Sometimes the same data must be entered or recorded multiple times, because the people generating the information cannot easily access the computer and enter it themselves.

Trained operators may spend hours daily entering commands and data -- a fatiguing task requiring concentration, typing proficiency, and innumerable keystrokes. Needless to say, fatigue reduces an operator's productivity and accuracy.

ENTERING DATA USING VOICE

Voice is probably the most natural, effective means of human communication. Compared to traditional keyboard input as a means to control and enter information into computers, voice can simplify the task, signficantly reduce the amount of training needed for computer operators, increase problem-solving speed and productivity, and reduce operator fatigue.

Voice input is already being used in these and other data entry applications:

- Manufacturing
  - Physical Inventory
  - Quality Inspections
  - Shipping & Receiving
- Medical
  - Laboratory Data Entry
  - Medical Reporting
- Distribution and Material Handling
  - Package Sorting and Routing
  - Inventory Management
  - Shipping and Receiving
- Financial Services
  - Insurance Form Completion
  - Financial Transactions
- Legal Services
  - Recording Time and Billing
- Order Entry
  - Large Volume Mail Order
  - Customer Service Account Status

Voice input can increase a business' capacity to handle ever-growing amounts of data more accurately and quickly. Timely information yields better customer service and better business decisions. In many cases, the payback period for adding voice input can be less than one year.

FASTER, MORE ACCURATE DATA ENTRY

Using voice, data entry operators can use their hands to hold an order form, skimming the form to find a certain field or box, and simply speak the information into the computer. More work gets done, even for data entry operators with excellent typing skills.

Voice input can virtually eliminate typographical errors, since translation of each spoken word is pre-defined and automatic.

Taking a medical example, it is more likely an operator could mistype "A116" for "A117", versus saying "heart" instead of "liver". In radiology, if one says "myocardial infarction", the KVS after recognizing the word will always spell it correctly.

Voice speeds data entry at Kodak's Health and Environmental Laboratory, where technicians use a KVS to report autopsy findings. They are unable to use the keyboard because their hands are gloved and dirty. With voice, the technicians no longer glance away from a slide to record data or look up code numbers in a manual.

NO NEED TO MEMORIZE MEANINGLESS CODES

Many data entry applications substitute shorter numerical codes to reduce the amount of typing. However, the operator is forced to translate these codes during the data entry process. This translation step takes valuable time, and can introduce errors.

At Kurzweil, shipping and receiving is done by voice; personnel simply state the product or part name, and the KVS converts the phrase into the proper code.

Voice input technology teaches machines to learn our language, rather than demanding that people learn computer language.

REDUCED USER FATIGUE, BETTER CONCENTRATION

All the advantages and efficiencies of voice input serve to reduce operator fatigue -- less typing, easy spoken commands, less distraction, faster response. Increased job satisfaction can reduce costly employee turnover.

SHORTER OPERATOR TRAINING PERIOD

Since voice input uses natural spoken words or phrases, even casual users can operate a voice-driven computer, without looking up commands in a manual. Typing ability is no longer a prerequisite. Given these factors , the KVS increases current staff's productivity, while expanding the available regional job pool to meet a business' data entry demands.

At Law School Data Admissions Services, training each new data entry person to enter transcript grades has been reduced from a week to just a few hours, using the KVS.

CAPTURE DATA AT ITS SOURCE

With voice, those skilled people who create or collect the information can enter it quickly and accurately, even if they're not proficient typists. They can make revisions and corrections as they enter the data verbally, without burdening support staff.

KURZWEIL VOICEPARTNERS™ Data Entry

Articulate Publications (*Medicalis, Dentalis*) - Medical/dental practice management software

Health Software (*The Medical Assistant, The Dental Assistant, Consent, The Electronic Tumor Registry*) - Medical/dental/hospital information software

Javelin Software (*Javelin*) - Integrated business management software

Program Resources (*Mindreader*) - Direct dictation word processing

Laboratory Technologies Corporation (*LABTECH NOTEBOOK, LABTECH CHROM*) - Scientific data acquisition, industrial process control

PRODUCT FAMILY

The Kurzweil Voicesystem™ (KVS), Kurzweil AI's first voice ecognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry applications.

The Kurzweil Voice Terminal™ (KVT), built from KVS technology, functions as an intelligent voice terminal that can be connected to large host systems without any modification to the host hardware or software.

The Kurzweil Voicewriter ™(KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

KURZWEIL VOICEPARTNER™ PROGRAM

The Kurzweil Voicepartner Program was established to provide a broad range of applications software from leading software suppliers. Kurzweil Voicepartners offer pre-defined core vocabularies to control their programs by voice, allowing users to implement their applications more quickly. In some cases, Voicepartners modify portions of their applications, such as the human interface, to further enhance the integration with voice input. Building upon the core vocabularies provided by the Voicepartner Program, users can modify the core vocabulary to meet their particular needs. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications.

KVS FEATURES

- **Large Vocabularies** - up to 1,000 user definable words or phrases can be recognized at a given time.
- **Real-time Response** - each word is displayed on the monitor as it is spoken.
- **Multiple Vocabularies** - can provide access to vocabularies for several applications each with up to 1,000 words or phrases.
- **Language Independent** - can be trained to understand any language.
- **High Accuracy and Reliability** - provides the level of accurate speech recognition necessary for commercial applications.
- **Spoken Commands** - a single spoken command can replace a complex string of keyboard strokes.
- **Simple Host Connection** - no modifications to existing host-resident applications are required.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of recognizing virtually any printed text and converting it to synthesized speech. In 1984 he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly recreate the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development and engineering. The company's technology leadership is backed by a commitment to full technical support and service. Kurzweil engineering and marketing staffs are working closely with customers to adapt voice recognition products to specific applications.

KURZWEIL AI

For further information on Kurzweil AI products contact your nearest sales office.

KURZWEIL APPLIED INTELLIGENCE, INC.

COMPANY HEADQUARTERS

411 Waverly Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverly Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Infomart, Suite 5001, 1950 Stemmons Freeway, Dallas, TX 75207, (214) 746-5023

West

N. Calif: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

International

Contact company headquarters for international distributor locations.

4/86 1.0

Handicapped:
Voice Opens the Workplace and the World to the Handicapped

For the physically handicapped, voice recognition offers real hope for bettering personal communications, education, and gainful employment. Voice is normally the most natural, effective form of communication. Until recently, voice recognition systems have been hampered by limited accuracy, insufficient vocabulary size, and high initial cost.

The Kurzweil Voicesystem™ (KVS) virtually eliminates these restrictions. For many people, voice input can overcome barriers in personal life, education, and work associated with physical handicaps. It can provide effective conversational communication with other people. In addition, voice can drive word processors, spreadsheets, database management packages, programming languages, and telecommunications systems.

Voice recognition technology can help the handicapped realize their full potential value to our society through productive employment.

FOR THE MOTOR-IMPAIRED

Using touch-driven computer input devices requires a degree of precision that persons with severe motor impairment may not possess. The KVS is already helping people afflicted with quadriplegia, cerebral palsy, and Parkinson's disease to communicate.

Computing without keyboarding via voice input can have a profound impact on the lives of the disabled. For the quadriplegic high school student injured in a football game, it can mean the resumption of his studies. For a quadriplegic programmer working for a bank it can mean increased productivity; he had employed a mouth stick to control the keyboard at a painstakingly slow pace. Using the KVS, he can program even faster than other programmers.

The Children's Hospital is using KVS speech recognition in a ground-breaking vocational program that will provide computer training and job placement for children and young adults whose severe physical handicaps prevent them from using standard keyboards, mice and other hand-held computer control devices. Dr.Howard Shane, Director of the four-year-old Communication Enhancement Clinic at Children's Hospital, envisions the pilot program as a vocational training center model that could be replicated nationwide in hospitals.

The Stanford Research Laboratory at Stanford University, with funding assistance provided by the Veterans Administration, has successfully developed a prototype of a voice-activated robot serving the physically disabled. With the KVS as the voice control module, the robot responds to voice commands to perform precise maneuvers like selecting a single flower from a bouquet in a vase, and bringing it to the robot operator.

FOR THE SPEECH IMPAIRED

Dr. Shane at Children's Hospital also plans to experiment with the KVS as a translation device. For people with impaired speech who still can make consistent sounds, including stroke victims and cerebral palsy patients, the KVS can act as a translation device when coupled to a voice output system. The person speaks into the microphone; the KVS recognizes sounds and then associates them with the proper words or commnds. The KVS can either send the recognized words to a display screen, or to a voice synthesizer. In this fashion, the KVS converts otherwise incomprehensible utterances to give people synthetic speech, which can be used for face-to-face and telephone conversation.

FOR THE BLIND

The KVS has been incorporated into a system with speech response which enables blind people to easily do word processing. Although many blind persons can use keyboards, the combination of speech input and speech response provides a powerful tool for interaction with computer systems.

FOR THE DEAF

The KVS can be used by people communicating with the deaf. The person communicating talks into the KVS, while the deaf person reads the speech that appears on a video screen. In this way, people unable or unwilling to learn sign language can converse with deaf people.

VALUED-ADDED RESELLERS - Handicapped:

Metropolitan Rehabilitation Services
Creating telemarketing job opportunities for the handicapped

Children's Hospital (Communications Enhancement Clinic)
Communications and vocational training/placement using voice-activated computers

Schlines Representation, Inc.
Manufacturers of complete voice workstation incorporating KVS input and synthetic speech output.

PRODUCT FAMILY

The Kurzweil Voicesystem™ (KVS), Kurzweil AI's first voice ecognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry applications.

The Kurzweil Voice Terminal™ (KVT), built from KVS technology, functions as an intelligent voice terminal that can be connected to large host systems without any modification to the host hardware or software.

The Kurzweil Voicewriter ™(KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

KURZWEIL VOICEPARTNER™ PROGRAM

The Kurzweil Voicepartner Program was established to provide a broad range of applications software from leading software suppliers. Kurzweil Voicepartners offer pre-defined core vocabularies to control their programs by voice, allowing users to implement their applications more quickly. In some cases, Voicepartners modify portions of their applications, such as the human interface, to further enhance the integration with voice input. Building upon the core vocabularies provided by the Voicepartner Program, users can modify the core vocabulary to meet their particular needs. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications.

KVS FEATURES

- **Large Vocabularies** - up to 1,000 user definable words or phrases can be recognized at a given time.
- **Real-time Response** - each word is displayed on the monitor as it is spoken.
- **Multiple Vocabularies** - can provide access to vocabularies for several applications each with up to 1,000 words or phrases.
- **Language Independent** - can be trained to understand any language.
- **High Accuracy and Reliability** - provides the level of accurate speech recognition necessary for commercial applications.
- **Spoken Commands** - a single spoken command can replace a complex string of keyboard strokes.
- **Simple Host Connection** - no modifications to existing host-resident applications are required.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of recognizing virtually any printed text and converting it to synthesized speech. In 1984 he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly recreate the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development and engineering. The company's technology leadership is backed by a commitment to full technical support and service. Kurzweil engineering and marketing staffs are working closely with customers to adapt voice recognition products to specific applications.

KURZWEIL AI

For further information on Kurzweil AI products contact your nearest sales office.

COMPANY HEADQUARTERS

411 Waverley Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverley Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Suite 161, 2915 L.B.J. Freeway, Dallas, TX 75234, (214) 247-3622

West

N. Calif.: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif.: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

International

Contact company headquarters for international distributor locations.

4/86 1.0

Health Care: Voice Input Improves Quality, Reduces Costs

Costs for health care in the United States currently exceed $400 billion. Health care professionals and system administrators are constantly pressured by the public and the government to provide better medical care. Still, they are expected to control spiraling medical costs. Since labor accounts for over 60% of the total operating expenses of a typical hospital, even small improvements in productivity can save millions of dollars.

VOICE INPUT FOR HEALTH CARE

Voice is probably the most natural, effective means of human communications, Compared to traditional keyboard/mouse input as a means to control and enter information into computers, voice can simplify the man/machine interface, significantly reduce the amount of training needed for computer operators, increase the accuracy and timeliness of health care data, and reduce operator fatigue.

Medical reporting

In medical disciplines like radiology, cardiology, and pathology, test information and results should be quickly and accurately processed. The reports must be delivered as soon as possible to physicians to make critical patient care decisions. Most medical reports are handwritten, or dictated to a transcriptionist, who is highly-skilled and in short supply in the health care industry. The process of dictation, transcription and proof reading can take from 6 hours to 3 days depending on the doctor's and transcriptionist's workload.

Using voice input, doctors can view test results and dictate diagnoses directly into the computer. They can use "voice macros" as labor-saving reporting techniques - for example, a single phrase like "broken femur" triggers a pre-programmed report entry describing the injury in precise medical terminology. The system quickly produces an accurate report for updating the patient's records, and distribution to the health care team. Voice dictation sidesteps the transcriptionist labor shortage, enables doctors to correct information during dictation, and speeds the medical reporting process.

Laboratory data entry

In pathology, pharmaceutical testing labs, and other areas of internal medicine, clinicians' hands and eyes are often too busy with microscopes, instruments, slides, blood samples, and equipment to type on a keyboard. Voice input allows them to dictate findings directly to the computer for reporting purposes, without requiring an extra staff person to take notes.

In the pharmacy

When filling prescriptions, pharmacists using voice input can record patient name and address, prescribing physician, dates, drug type, dosages, and other information for records and billing. The pharmacy computer can automatically record and process this data, making the pharmacy more productive.

Direct rapid access to databases

Health care professionals who need quick access to patient or administrative records stored in a central database do not want to master typing or computer programming to obtain this information. For example, a pharmacist may want to compare a patient's record with drug database information before administering dosage; a physician with an emergency case may need a file in seconds. Voice recognition helps these professionals retrieve the data when needed, without resorting to the clerical staff or data-entry operators.

Medical practice management

There is a wide variety of PC-based software available to assist medical practices and hospitals with critical information management. Medical practice management systems can help physicians, dentists, anesthesiologists, radiologists, veterinarians, chiropractors and other health care professionals maintain patient and practice records, including diagnostic information, mailing lists, responsible party information, insurance IDs and codes, and billing/payment records. These and other software packages can be controlled by voice.

Patient examination and diagnoses

Here again, voice dictation to the patient's computer-stored record enables the physician to concentrate on the examination while verbally recording case data. Keyboarding or note-taking is not practical during certain exam procedures routinely performed in areas such as ophthamology and gastroenterology.

KURZWEIL VOICEPARTNERS™ - Health Care

Articulate Publications (*Medicalis, Dentalis*)-Practice management software

Health Software (*The Medical Assistant, The Dental Assistant, Consent, The Electronic Tumor Registry*)-Medical information software

PRODUCT FAMILY

The Kurzweil Voicesystem™ (KVS), Kurzweil AI's first voice ecognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry applications.

The Kurzweil Voice Terminal™ (KVT), built from KVS technology, functions as an intelligent voice terminal that can be connected to large host systems without any modification to the host hardware or software.

The Kurzweil Voicewriter ™(KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

KURZWEIL VOICEPARTNER™ PROGRAM

The Kurzweil Voicepartner Program was established to provide a broad range of applications software from leading software suppliers. Kurzweil Voicepartners offer pre-defined core vocabularies to control their programs by voice, allowing users to implement their applications more quickly. In some cases, Voicepartners modify portions of their applications, such as the human interface, to further enhance the integration with voice input. Building upon the core vocabularies provided by the Voicepartner Program, users can modify the core vocabulary to meet their particular needs. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications.

KVS FEATURES

- **Large Vocabularies** - up to 1,000 user definable words or phrases can be recognized at a given time.
- **Real-time Response** - each word is displayed on the monitor as it is spoken.
- **Multiple Vocabularies** - can provide access to vocabularies for several applications each with up to 1,000 words or phrases.
- **Language Independent** - can be trained to understand any language.
- **High Accuracy and Reliability** - provides the level of accurate speech recognition necessary for commercial applications.
- **Spoken Commands** - a single spoken command can replace a complex string of keyboard strokes.
- **Simple Host Connection** - no modifications to existing host-resident applications are required.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of recognizing virtually any printed text and converting it to synthesized speech. In 1984 he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly recreate the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development and engineering. The company's technology leadership is backed by a commitment to full technical support and service. Kurzweil engineering and marketing staffs are working closely with customers to adapt voice recognition products to specific applications.

KURZWEIL AI

For further information on Kurzweil AI products contact your nearest sales office.

KURZWEIL APPLIED INTELLIGENCE, INC.

COMPANY HEADQUARTERS

411 Waverly Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverly Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Infomart, Suite 5001, 1950 Stemmons Freeway, Dallas, TX 75207, (214) 746-5023

West

N. Calif: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

International

Contact company headquarters for international distributor locations.

4/86 1.0

Manufacturing:
Voice a Critical Link for Computer Integrated Manufacturing

Manufacturers are rapidly moving towards the goal of Computer Integrated Manufacturing (CIM). To implement CIM, computer systems involved in the every phase of the production cycle must be linked and integrated to allow timely sharing of data. Data must be entered accurately and quickly. Once stored in the database, this information must be easily accessible to decision makers for analysis. And finally, production machinery must be efficiently controlled and monitored.

Since billions of dollars have already been invested in manufactuirng plants, equipment, and personnel, productivity increases of just a few percent can save millions of dollars.

Voice is a key towards realizing the goals of computer-integrated manufacturing.

VOICE DATA ENTRY

Managing a volume manufacturing operation typically requires entering massive amounts of data, especially in areas such as quality control, order entry, materials management, and shipping and receiving. Often, the operator must interrupt the task each time he shifts his eyes or moves his hands to the keyboard. Voice input saves valuable time since hands and eyes are now free for ther tasks.

Order Forms

For example, an employee may be manually sorting through order forms, while entering data from a certain field. Using voice, the operator avoids the constant interruption and fatigue caused by moving between the form, the keyboard, and the display screen.

Voice input improves accuracy. Each spoken word, once recognized, is automatically translated into text thereby eliminating typographical errors. Using english words rather than computer-based codes for data entry and commands, mistakes are more likely to be identified and corrected during the data entry process.

The person who collects the data can enter it directly.The elimination of the intermediate transcription step, not only saves time, but allows the data to be checked interactively, as it is entered, thereby reducing errors. Furthermore, since typing and computer knowledge are not required, the labor pool for data entry operators is increased dramatically.

Inspection and Test

Voice produces similar benefits while performing inspection and test of printed circuit boards. Both hands can be used to handle the board or test probe, while eyes inspect for defects, and voice enters the data into the computer system.

Physical Inventory

Physical inventory also can benfit from voice input. Parts can be entered by verbal description, such as "power supply transformer" rather than typing "02-08733-34". Additional data such as quantity, weight, and physical condition can easily be entered simply by speaking. Wireless microphones and voice output can be combined with voice recognition to develop systems for mobile use throughout the factory.

In each of these examples, with voice, work flow is not interrupted so jobs are done faster and more accurately.

INFORMATION ACCESS FOR DECISION MAKERS

Through simple spoken commands, professionals can get the information they need from the computer, when they need it. Access to information is no longer limited to skilled typists or computer operators. Faster access to data means better decisions, and reduced demand for support staff translates into savings of both time and money.

With voice input, there is no need to remember and type complex codes. A single voice command or "voice macro" can replace hundreds of equivalent keystrokes, and can be easily adapted to meet your particular needs. Speaking a command such as "Print MRP" is much simpler than remembering and typing a long sequence of keystrokes to generate and format the report.

Voice recognition technology makes computers learn our language instead of our having to learn the computer's language. Voice input can be the tool that makes the computers truly usable by middle and upper management, and in this way contributes to the goal of CIM.

MANUFACTURING COMMAND AND CONTROL

Controlling machines usually requires that the operator enters commands from a fixed control station. The operator must learn the proper codes and must always have easy access to the keyboard.

With voice input, anoperator can control the machine even at a distance from the keyboard. Hands and eyes are free to perform an inspection or repair task, while voice simultaneously controls the machine. In some manufacturing plants, preventing even a single machine shut-down can easily justify the investment for voice input.

An emerging application has been to couple voice input with robotics. This flexible production tool can easily be controlled by operators without extensive training.

PRODUCT FAMILY

The Kurzweil Voicesystem™ (KVS), Kurzweil AI's first voice ecognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry applications.

The Kurzweil Voice Terminal™ (KVT), built from KVS technology, functions as an intelligent voice terminal that can be connected to large host systems without any modification to the host hardware or software.

The Kurzweil Voicewriter ™(KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

KURZWEIL VOICEPARTNER™ PROGRAM

The Kurzweil Voicepartner Program was established to provide a broad range of applications software from leading software suppliers. Kurzweil Voicepartners offer pre-defined core vocabularies to control their programs by voice, allowing users to implement their applications more quickly. In some cases, Voicepartners modify portions of their applications, such as the human interface, to further enhance the integration with voice input. Building upon the core vocabularies provided by the Voicepartner Program, users can modify the core vocabulary to meet their particular needs. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications.

KVS FEATURES

- **Large Vocabularies** - up to 1,000 user definable words or phrases can be recognized at a given time.
- **Real-time Response** - each word is displayed on the monitor as it is spoken.
- **Multiple Vocabularies** - can provide access to vocabularies for several applications each with up to 1,000 words or phrases.
- **Language Independent** - can be trained to understand any language.
- **High Accuracy and Reliability** - provides the level of accurate speech recognition necessary for commercial applications.
- **Spoken Commands** - a single spoken command can replace a complex string of keyboard strokes.
- **Simple Host Connection** - no modifications to existing host-resident applications are required.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of recognizing virtually any printed text and converting it to synthesized speech. In 1984 he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly recreate the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development and engineering. The company's technology leadership is backed by a commitment to full technical support and service. Kurzweil engineering and marketing staffs are working closely with customers to adapt voice recognition products to specific applications.

KURZWEIL AI

For further information on Kurzweil AI products contact your nearest sales office.

KURZWEIL APPLIED INTELLIGENCE, INC.

COMPANY HEADQUARTERS

411 Waverly Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverly Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Infomart, Suite 5001, 1950 Stemmons Freeway, Dallas, TX 75207, (214) 746-5023

West

N. Calif: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

International

Contact company headquarters for international distributor locations.

4/86 1.0

Technical Workstations:
Voice-Based Technical Workstations Boost Productivity

Workstations make problem-solving, design, and testing easier and quicker for engineers, scientists, architects, technical writers, and designers. However, most workstations still require mastery of complex commands for most effective usage. Once trained, workstation operators spend hours daily entering commands and data -- a fatiguing task requiring concentration, typing proficiency, and innumerable keystrokes. The constant refocusing needed to operate keyboards and mice can break a user's concentration. This fatigue reduces productivity and accuracy.

VOICE COMMAND AND CONTROL

Voice is probably the most natural, effective means of human communication. Compared to traditional keyboard/mouse input as a means to control and enter information into computers, voice can simplify the man/machine interface, significantly reduce the amount of training needed for computer operators, increase problem-solving speed and productivity, and reduce operator fatigue.

Voice input is being used to complement keyboard and mouse input in these technical workstation applications:

- Computer-aided Design (CAD)
- Computer-aided Engineering (CAE)
- Scientific and Industrial Data Acquisition
- Software Programming and Debugging
- Electronic Printing and Publishing

Voice input is easier, faster, and more accurate than mouse or keyboard entry, especially for menu selection. Productivity gains from 20-30% have been measured, derived primarily from better concentration on the task.

All these benefits contribute to a quicker return on investment for workstation hardware, software, service, and personnel. In many cases, the payback period for adding a speech-recognition device can be less than one year.

BETTER CONCENTRATION ON THE DESIGN OR TASK

Without the distraction of remembering and typing commands with complex codes and syntaxes, or searching through hierarchical menus, operators can concentrate better on their projects. One simple voice command -- a "voice macro" -- can replace a series (sometimes hundreds) of keystrokes.

The handling of symbol, part, and signal names becomes easier. For example, an electronic design engineer using FutureNet can say "Load 3-to-8 decoder", rather than typing "*138A (RET)". The designer could use a single spoken phrase -- "Run simulation" -- that automatically loads the necessary files, sets input conditions, runs a test, and prints out the results.

FASTER, MORE ACCURATE DATA ENTRY

Voice input can reduce typographical errors, since translation of each spoken command is pre-defined for the application. Gateway Design Automation Corporation, manufacturer of VERILOG™ circuit design simulation software, uses voice input to replace standard VERILOG operating commands with their own spoken command words, virtually eliminating errors from typing individual commands like "$" and ";".

The operator simply looks at the display and says a command word -- for example, "Zoom" -- which bypasses menus and keyboard input. The operator does not have to continually switch between the keyboard, screen menus, and the mouse, thus freeing valuable time for creative thinking.

For nearly a year at Kurzweil AI, engineers have been designing circuit boards with voice-driven DASH™ workstations from FutureNet. They report better concentration, and a 30% productivity improvement. In fact, some found that after using the voice system for several days, they had difficulty returning to a standard DASH workstation.

REDUCED USER FATIGUE

Operators use workstations interactively for hours at a time, without a support person to enter prepared data. The concentration and typing can be tiresome, especially for professionals and technicians paid for their expertise and creativity. All the advantages of voice input also serve to reduce operator fatigue -- less typing, easy spoken commands, less distraction, faster response.

SHORTER OPERATOR TRAINING PERIOD

Since voice input uses natural spoken words or phrases, even casual users can operate a voice-driven computer, without looking up commands in a manual. Valuable professionals can quickly learn to use the workstation to its full capability, and wider group of people can use the workstation for their tasks.

FASTER TIME TO MARKET FOR NEW PRODUCTS

Time that voice input saves on crucial design and technical projects can have significant benefits. Even a few days saved can mean thousands of extra dollars in sales revenue and profits generated on successful new commercial products. For scientists, it can mean completing more projects within current research budgets.

KURZWEIL VOICEPARTNERS™ Technical Workstations

Laboratory Technologies Corporation (*LABTECH NOTEBOOK, LABTECH CHROM*) - Scientific data aquisition, industrial process control.

FutureNet Corporation (*DASH*) - Circuit schematic capture and logic design.

Gateway Design Automation Corporation (*Verilog*) - Circuit design simulation and testing

VALUE-ADDED RESELLERS -- Technical Workstations:

Manufacturing Consulting Services (*Anvil*) -Mechanical CAD Software

PRODUCT FAMILY

The Kurzweil Voicesystem™ (KVS), Kurzweil AI's first voice ecognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry applications.

The Kurzweil Voice Terminal™ (KVT), built from KVS technology, functions as an intelligent voice terminal that can be connected to large host systems without any modification to the host hardware or software.

The Kurzweil Voicewriter ™(KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

KURZWEIL VOICEPARTNER™ PROGRAM

The Kurzweil Voicepartner Program was established to provide a broad range of applications software from leading software suppliers. Kurzweil Voicepartners offer pre-defined core vocabularies to control their programs by voice, allowing users to implement their applications more quickly. In some cases, Voicepartners modify portions of their applications, such as the human interface, to further enhance the integration with voice input. Building upon the core vocabularies provided by the Voicepartner Program, users can modify the core vocabulary to meet their particular needs. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications.

KVS FEATURES

- **Large Vocabularies** - up to 1,000 user definable words or phrases can be recognized at a given time.
- **Real-time Response** - each word is displayed on the monitor as it is spoken.
- **Multiple Vocabularies** - can provide access to vocabularies for several applications each with up to 1,000 words or phrases.
- **Language Independent** - can be trained to understand any language.
- **High Accuracy and Reliability** - provides the level of accurate speech recognition necessary for commercial applications.
- **Spoken Commands** - a single spoken command can replace a complex string of keyboard strokes.
- **Simple Host Connection** - no modifications to existing host-resident applications are required.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of recognizing virtually any printed text and converting it to synthesized speech. In 1984 he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly recreate the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development and engineering. The company's technology leadership is backed by a commitment to full technical support and service. Kurzweil engineering and marketing staffs are working closely with customers to adapt voice recognition products to specific applications.

KURZWEIL AI

For further information on Kurzweil AI products contact your nearest sales office.

KURZWEIL APPLIED INTELLIGENCE, INC.

COMPANY HEADQUARTERS

411 Waverly Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverly Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Infomart, Suite 5001, 1950 Stemmons Freeway, Dallas, TX 75207, (214) 746-5023

West

N. Calif: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

International

Contact company headquarters for international distributor locations.

4/86 1.0

Natural Language: Decision Makers Can Talk to Their Databases in Plain English

The thousands of currently available databases contain information with great potential to improve decisions in business, research, health care, and government. Billions of dollars have already been invested in the computer equipment, software, data and personnel. To unlock the true potential of these databases, the challenge has been to first organize the information in a useful format, and then to make the information readily accessible to those people who need it most.

Natural language interfaces, which invite the user to ask database questions in plain, conversational English, have been a boon to knowledge workers who are not necessarily programmers. These interfaces, however, require typing skills beyond the capability of many middle and upper level managers.

VOICE LETS YOU THINK ALOUD

Voice is probably the most natural, effective means of human communication. Compared to traditional keyboard input as a means to control computers, voice simplifies the man/machine interface, significantly reduces the amount of user training needed, increases problem-solving speed and productivity, and reduces user's fatigue.

Phrasing questions for a database is easy and direct when the user can "think aloud". For example, users of Artificial Intelligence Corporation's INTELLECT can simply ask for the comparison of planned and actual costs for the last quarter. The user's spoken request is recognized directly by the system, and the correct data appears as the words are spoken.

INFORMATION ACCESS FOR DECISION MAKERS

Through simple spoken commands, professionals can get the information they need from the database, when they need it. Access to information is no longer limited to skilled typists or computer operators.

Voice gives decision makers interactive access to information. For example, a sales manager may have originally requested sales revenue by district. Noticing that one district was ahead of plan, the manager may now ask for a more detailed report by sales representative and major account.

A simple change in the spoken request can place the desired information within reach without delay.

Faster access to data means better decisions, and reduced demand for support staff translates into savings of both time and money.

VOICE IS FASTER, MORE ACCURATE

Database inquiry using voice can be more than three times faster than typing the same request.

"Voice macros" can be defined by the user to replace a series (sometimes hundreds) of keystrokes. In addition, voice can reduce typographical errors, since translation of each spoken word is pre-defined for the application.

One study showed that adding voice to Microrim's Clout cut the elapsed session time by 30%. Such savings mean increased productivity.

SHORTER TRAINING PERIOD

Since voice input uses natural spoken words or phrases, even casual users can operate a voice-driven database without looking up commands in a manual, or typing long phrases.

Valuable professionals can quickly learn to use the database to its full capability, and a wider group of people can access information for their tasks.

KURZWEIL VOICEPARTNERS™ Natural Language

Alexander-Robertson, Ltd. (*Karl*) - Natural language personal filing system.

Artificial Intelligence Corporation (*INTELLECT*) - Natural language access to mainframe databases.

Brodie Associates (*Language Workbench*) - Programmer's toolkit for developing natural language interfaces to applications software.

Finder Information Tools (*Finder: The Retrieval Software*) - Database management and retrieval software

Microrim (*Clout*) - Natural language interface to R:base 4000 and 5000 database programs.

Program Resources, Inc. (*Mindreader*) - Intelligent menu-driven text entry based on context.

PRODUCT FAMILY

The Kurzweil Voicesystem™ (KVS), Kurzweil AI's first voice ecognition product with vocabularies of up to 1,000 words, has been in production since 1985. The KVS has been incorporated into diverse command-and-control and data entry applications.

The Kurzweil Voice Terminal™ (KVT), built from KVS technology, functions as an intelligent voice terminal that can be connected to large host systems without any modification to the host hardware or software.

The Kurzweil Voicewriter ™(KVW) is designed to provide dictation and word processing with an expanded vocabulary, enabling general purpose text creation.

The Kurzweil Telerecognizer™ (KTR) is being designed for speech recognition of limited vocabularies over telephone lines, with audio response and speaker independence.

Kurzweil AI's commitment to technological leadership includes significant on-going investments to commercialize voice recognition products with continuous speech recognition, speaker independence, and larger vocabularies, while lowering costs over time.

KURZWEIL VOICEPARTNER™ PROGRAM

The Kurzweil Voicepartner Program was established to provide a broad range of applications software from leading software suppliers. Kurzweil Voicepartners offer pre-defined core vocabularies to control their programs by voice, allowing users to implement their applications more quickly. In some cases, Voicepartners modify portions of their applications, such as the human interface, to further enhance the integration with voice input. Building upon the core vocabularies provided by the Voicepartner Program, users can modify the core vocabulary to meet their particular needs. The list of Kurzweil Voicepartners has grown rapidly as software manufacturers have been made aware of the benefits of voice input for their applications.

KVS FEATURES

- **Large Vocabularies** - up to 1,000 user definable words or phrases can be recognized at a given time.
- **Real-time Response** - each word is displayed on the monitor as it is spoken.
- **Multiple Vocabularies** - can provide access to vocabularies for several applications each with up to 1,000 words or phrases.
- **Language Independent** - can be trained to understand any language.
- **High Accuracy and Reliability** - provides the level of accurate speech recognition necessary for commercial applications.
- **Spoken Commands** - a single spoken command can replace a complex string of keyboard strokes.
- **Simple Host Connection** - no modifications to existing host-resident applications are required.

COMPANY PROFILE

Kurzweil AI was formed specifically to tackle the problem of large vocabulary speech recognition. Its founder, Raymond Kurzweil, is an established expert in pattern recognition techniques. In 1976, he introduced the Kurzweil Reading Machine for the blind, the first optical scanning system capable of recognizing virtually any printed text and converting it to synthesized speech. In 1984 he introduced the Kurzweil 250, the first keyboard synthesizer to convincingly recreate the sounds of orchestral instruments.

Kurzweil AI has recruited an interdisciplinary team of specialists from the fields of acoustic science, linguistics, signal processing, software development and engineering. The company's technology leadership is backed by a commitment to full technical support and service. Kurzweil engineering and marketing staffs are working closely with customers to adapt voice recognition products to specific applications.

KURZWEIL AI

For further information on Kurzweil AI products contact your nearest sales office.

KURZWEIL APPLIED INTELLIGENCE, INC.

COMPANY HEADQUARTERS

411 Waverly Oaks Road
Waltham, MA 02154
Tel: (617) 893-5151
TELEX: 9103809408

SALES OFFICES

East

Boston: 411 Waverly Oaks Road, Waltham, MA 02154, (617) 893-5151

New York: 120 Wood Avenue South, Iselin, NJ 08830, (201) 321-0555

Wash. D.C.: 6411 Ivy Lane, Greenbelt, MD 20770, (301) 982-0666

Central

Chicago: Suite 400, 8410 W. Bryn Mawr Avenue, Chicago, IL 60631, (312) 380-1215

Dallas: Infomart, Suite 5001, 1950 Stemmons Freeway, Dallas, TX 75207, (214) 746-5023

West

N. Calif: Suite 200, 1301 Shoreway Road, Belmont, CA 94002, (415) 591-1708

S. Calif: Suite 1100, 2049 Century Park East, Los Angeles, CA 90067, (213) 203-9909

International

Contact company headquarters for international distributor locations.

4/86 1.0

THE TECHNOLOGY OF THE KURZWEIL VOICE WRITER

BY RAYMOND KURZWEIL

The present office system provides a clue to future applications for the deaf

Editor's note: This article is not a review of the KVW; it is a look at a technology that may be available on personal computers in the future.

THE KURZWEIL VOICE WRITER (KVW) is a voice-activated word processor with a relatively unrestricted user-specific vocabulary. The system starts with a vocabulary of at least 5000 frequently used words in the English language. It subsequently adds the words you use that are not part of its initial vocabulary and eventually deletes those words that you never use. Total vocabulary, depending on the KVM model, will be in the 7500- to 20,000-word range.

Voice is our most effective and rapid means of communication, and the ability to interact with computerized information services and devices by voice, without the restrictions of artificial vocabularies or syntax, is expected to be of major benefit. The primary application of the KVW is to automate the creation of written text, which is a fundamental activity in the office. Combining large-vocabulary ASR (automatic speech recognition) with natural-language understanding would also enable professionals and executives to make inquiries of database-management systems or management information systems verbally in natural language instead of through a keyboard.

One planned application of this technology is to create a speaker-independent version of the KVW to serve as a display telephone for the deaf. This would enable a deaf person to hold a phone conversation without being restricted to speaking to other deaf people who have compatible TDDs (telecommunications device for the deaf). It is not yet available but the technology that will be used in its creation is described in essence in this article.

The KVW as it currently exists requires only that you can speak and that you can see. Motion and hearing impairments are not obstacles in its operation. The current version of the KVW is for the business community, but it fills a need for many disabled persons as well. The initial KVM model, which can be shared by multiple users (one at a time) is expected to be introduced this year at a price under $20,000. Future models of both single-user and multiuser systems are expected to be in the $4000 to $10,000 range. While this is beyond the price range of most individuals, the technology is the clue to future, more individually affordable solutions.

LARGE-VOCABULARY ASR

There are two difficulties involved in creating large-vocabulary ASR. First, you must create a set of linguistic and speech-recognition algorithms that provide the requisite recognition power and that are capable of resolving the fine distinctions and ambiguities that are inevitable when you deal with a large natural vocabulary. The incidence of "perplex clusters" (words that differ by only one phonetic feature) is much higher for

(continued)

Raymond Kurzweil is the founder and chairman of Kurzweil Computer Products, Kurzweil Music Systems Inc., and Kurzweil Applied Intelligence Inc. He received a B.S. degree from MIT and an honorary Ph.D. from Hofstra University. He can be reached at 411 Waverly Oaks Rd., Waltham, MA 02154.

To take maximum advantage of a multiple-expert strategy, you must combine the results from each expert in a way that recognizes its unique strengths and weaknesses. In general, the system can quickly and accurately resolve each recognition within a small perplex set of words. After this initial cut of the vocabulary to a small set (ranging from one word to a few dozen), the expert-management techniques depend to a great extent on the nature of the resulting perplex set. Some of the expert-management techniques are knowledge-based. For example, the handling of homonym sets is done through a single expert that is capable of differentiating between homonyms based on context. Other techniques involve probability: the methods of combining the probabilities from each expert are controlled by statistics on how the various experts have performed for different types of perplex sets. Some of these parameters are derived from statistics gathered during a particular user's time on the system and thus form part of the overall user-adaptation process.

Language Experts

A number of experts try to predict the likelihood of different words occurring at a particular lexical entry point based entirely on context. These experts use a variety of information-theory as well as sentence-parsing techniques.

The sentence-parsing expert is similar to the type of parser used in some natural-language understanding programs in that a tree-like structure is generated showing the part of speech of each word and its relationship to other words in the sentence. One significant difference is that the KVW parser is able to generate parses on incomplete sentences. At a particular point in a dictated sentence, we have only the "left" part of the sentence (from the beginning up through and not including the current word). Based on each parse of the incomplete sentences as they come in, the parsing expert is able to assign probabilities to different parts of speech. Rather than the eight or nine basic parts of speech that grade school children are familiar with (noun, verb, adjective, etc.), the KVW parser uses approximately 200 types representing subcategories of the basic parts of speech. This degree of specificity enables the parsing expert to increase the value of its predictions. Using a lexicon of approximately 50,000 words that indicates the likelihood of different parts of speech for each word, the parsing expert is able to assess the likelihood of different words. In particular, the parsing expert is good at eliminating choices that are syntactically unlikely.

There is a fortunate orthogonality between the strengths of the acoustic experts and those of the language experts. For example, most homonyms represent significantly different syntactic types that can be determined from context. "Two," "to," and "too" represent very different grammatical categories with readily identifiable word contexts. Also, short function words, which tend to be more difficult for an acoustic recognizer, are actually easier for the language model to make predictions for.

Acoustic Experts

The acoustic experts share an acoustic front-end processor that includes a high-resolution digitization (over 96-decibel dynamic range) and a robust filter bank made up of several hundred two-pole filter elements with 24-bit accuracy. The resulting spectral data is subsequently processed through a series of normalizations and other transformations to reduce variability and preserve feature invariance. Some of the transformations are based on an auditory model similar in many ways to the human ear's auditory front-end processing.

The acoustic experts utilize a RAM storage of word models, which are updated after each utterance. The acoustic experts are capable of evaluating the likelihood of every word model for a given test token, although the expert manager may request that a particular acoustic expert evaluate only a subset of the models based on the results of earlier experts.

Parallel-Processing Architecture

One area that uses extensive parallel processing is the front-end filtering. In order to make the fine distinctions necessary to handle the perplexity of a large vocabulary, a great deal of accuracy and resolution is needed in the number of filter channels and the accuracy of both the sample stream and the filters. Filtering is handled by the KSC2408 filter chip (from Kurzweil Semiconductor, a division of Kurzweil Applied Intelligence Inc.) with several transformations are handled by multiple conventional microprocessors.

Using the KVW

In dictation mode, you simply speak your text in a rapid, discrete manner, with brief pauses between words. The pause required between words is adjustable and should be set just long enough to reduce or eliminate the ambiguity between word pauses and stopgaps within a word. In general, this figure ranges from 100 to 250 milliseconds. The system responds within 500 ms after the end of each word by displaying the recognized word on the screen. A special status line displays any alternate word choices. In trials of the KVW, when the system has chosen the wrong word, the correct word has usually been the first or second alternate given.

The basic mode of operation is to speak into the system and watch the two-pole filters used for each filter channel. Implementing the 2408's filter algorithm (for a single two-pole filter) on a 68000 requires five seconds to process one second of speech, or five times real time. Each KSC2408 chip includes eight such filters (which operate in real time) and is thus equivalent to forty 68000 microprocessors (for the 2408 filter algorithm). The current model 1 KVW uses 25 KSC2408 chips, which is equivalent to using a thousand 68000 microprocessors for the filtering operation.

The equivalent of several thousand additional 68000 microprocessors (for certain dedicated algorithms, not for general-purpose computation) is provided by other special circuits used in the acoustic-matching process. The language experts and elements of the acoustic-recognition process such as normalization and other text appear. You don't need to be aware of what is in the active vocabulary. You simply speak and let the vocabulary-adaptation process proceed automatically.

You can also enter commands by voice. To distinguish commands from text, you enter a command mode either by depressing a function key or by speaking an appropriate unique verbal "Enter command mode" instruction (for example, "blix"). Once you enter command mode, you can switch among different types of commands to go, for example, from application-program commands to operating-system commands.

The primary mode of integrating the KVW's capabilities with an application program is through "transparent" integration. In this mode, the KVW simulates the keyboard. Recognized text and commands are converted

into appropriate character strings and transmitted to the operating system as if they came from the keyboard. The character strings come in through a special serial line and an appropriate driver intercepts them and presents them to the operating system as having come from the keyboard.

User Interface

One user interface that has been proposed for ease of use includes a pointing device (such as a mouse) to control the cursor, which is not easily manipulated by either keystrokes or verbal commands. The mouse would have two buttons, one to toggle between text and command mode and the second to correct errors. Again, you would have the choice of using these two buttons or using verbal commands. You would have relatively little use for the keyboard. Being able to correct most errors, go back and forth between text and commands, and control the location of the cursor would provide most of the control necessary aside from the actual verbal dictation of the text and commands.

To take this concept one step further, you could combine a flat-panel display with a touch-sensitive surface to provide a "pad" that you would hold in your lap or on your desk. As you speak to the pad, words would appear on its surface display. To control the cursor for insertion, deletion, or replacement operations, you would simply point to the screen. The two basic functions of error correction and toggle-to-command mode would be provided by either displayed "buttons," real buttons, or voice command (at your option). For the occasional requirement to type in a new vocabulary word, a QWERTY keyboard could be displayed on the screen.

Physical Configuration

The KVW consists of an approximately 100-megabyte Winchester disk, four circuit boards, and a power supply in a standard rack-mountable cabinet. While it would be possible to sit the KVW server next to the workstation it serves, it is generally found in a separate location. Thus, you interact only with your workstation and a microphone. The microphone can be either head-mounted, worn on your lapel, or desk-mounted. It is connected to a small box that digitizes the signal and transmits it on a high-speed serial line.

Future Directions

Future applications of the KVW technology include integration with natural-language-understanding systems, domain-specific expert systems, text-to-speech synthesizers, and a variety of application packages to provide executive assistants that are powerful and easy to use. Such systems will have access to the internal databases and MIS (management information system) information of the user's own organization as well as public, semipublic, and restricted-access databases accessed by telecommunications. Professionals, executives, students, and others will be able to converse with such systems to conduct rapid research and inquiry into a variety of questions of interest. Such questions might involve information retrieval ("How did the sales in our Western region for the past quarter compare to those of our three largest competitors?") as well as substantive analysis ("Which financing option for the proposed capital acquisition is best supported by our current balance sheet?"). Questions would be asked by voice in natural language. The questions would be clarified through two-way voice communication (or display), and final answers would be provided by either voice, display, or printout, as appropriate.

The acoustic experts in the KVW are adaptable to continuous speech input. The computation requirements must be increased to handle connected speech, as must the recognition power requirements to handle the additional perplexity of word segmentation, interword coarticulation, and function word reduction. It is expected that economically viable systems that can handle continuous speech will follow discrete-word KVWs within a few years.

The KVW techniques are also adaptable to European languages. The acoustic experts require very little change. The principal changes necessary to the language experts are (1) to provide the appropriate grammar rules to the parsing expert (although the parsing-expert algorithms themselves don't require substantial change) and (2) to train the language experts on appropriate foreign-language text. Foreign-language KVWs will probably follow the English KVW within a few years.

Handling Japanese requires more work than do European languages such as French or German. While Japanese has only about 120 syllables (compared to around 10,000 in English), the syllable set is a perplex one, with many syllable pairs being distinguished only by the duration of the vowel. Also, the differences in Japanese syntax require modifying more than just the parsing expert's grammar rules. Most of the KVW's techniques are, however, appropriate to the language, and a Japanese machine is feasible.

A number of configurations of a speech-to-display sensory aid for the deaf using the KVW technology have been proposed, which the company plans to pursue. Alternatives range from a speaker-independent version of the KVW (with an increased error rate) to a system that displays phonetic transcriptions rather than words. Such a phonetic transcription would contain some insertion, deletion, and substitution errors but could be understood by the user with appropriate training.

Conclusion

The introduction of large-vocabulary ASR is expected to provide dramatic productivity gains in creating written text, an optimal mode of communication between persons and intelligent computerized devices and services for information retrieval and analysis, as well as improved understanding and communication for the deaf population. ■

Kurzweil Applied Intelligence Incorporated
411 Waverley Oaks Road
Waltham, MA 02154-8465
617-893-5151 Telex: 9103809408

KAI #65•00010•000