Automated Speech Technology for Gastrointestinal Endoscopy Reporting and Image Recording

Oliver W. Cass, M.D., F.A.C.P.,
Assistant Professor of Medicine, University of Minnesota, Minneapolis

Abstract

Endoscopic databases will not be used until they are more efficient than (but as easy for physicians to use as) traditional dictate-and-transcribe methods. VoiceGI, a voice recognition system for endoscopic data entry, editing, storage and retrieval, and for capturing and storage of endoscopic images, is efficient and easy to use.

1. Proposed Demonstration

We will demonstrate VoiceGI, an integrated software system using speaker-independent voice recognition or menus for data entry, editing, storage, and retrieval of endoscopic reports and for capturing and storage of endoscopic images. There will be a simulated endoscopic examination, during which the endoscopist will capture and save digital endoscopic images using voice commands. Then, the endoscopist will produce a report by voice. Paper copies will be immediately produced by a laser printer for the patient's chart and for the referring physician (cover letter included). The data in the report will be saved in dBASE IV in a format compatible with the American Society for Gastrointestinal Endoscopy (ASGE) lexicon and the Pentax Reporter, a menu-driven endoscopic report generator.

Terms not in the ASGE lexicon may be used without additional "training" (a procedure for adding new vocabulary to the system) because a "free-text" vocabulary of over 3000 words was drawn from retrospective analysis of over 2000 endoscopic reports. New words will be trained by the endoscopist during report preparation. Editing, report and patient recall (recheck of a gastric ulcer at three months), statistics (number of endoscopies/ month), and quality assurance data (monthly complications) will also be demonstrated.

This program uses the Kurzweil VoiceMED(tm) software for recognition of 5000-20,000 words. The structure of the report was designed with the Kurzweil Knowledge Base Editor, which will also be demonstrated.

2. Relevance to Medical Informatics and Medical Care

Endoscopic reports can be produced by handwriting, filling out a form, dictation, or computer. Endoscopic data can be entered by keyboard into fill-in-the-blank forms on a computer screen using arrow keys or mouse to select one choice from a menu, or by speaking words or phrases into a device that recognizes the sound waves and associates them with text (voice recognition).

Handwriting produces an immediate record, but is often illegible and time consuming. Fill-in-the-blank forms take less time and give immediate results, but are often illegible and inflexible. Almost all endoscopists currently use dictation, which is transcribed, returned to the physician for editing, often retyped, again reviewed, signed, and sent to the chart and referring physician. The procedure is expensive and time consuming and does not provide for standardization of terms, rapid verification of accuracy, rapid documentation of findings in the patient's chart, immediate communication to referring physicians, association of verbal data with endoscopic pictures, easy recall of patient data for follow-up examinations, comparison of studies done at multiple institutions, preparation of statistics (locally and nationally), documentation for endoscopist or third-party payor certification, or quality assurance. Because of the lag time inherent in the process, a handwritten note is typically placed in the patient's chart.

The first computerized programs for endoscopic data entry are merely computer presentations of forms in use for the past 20 years. To fill in the blanks on the computer screen, physicians who use these programs must know how to type. An alternative program (The Endoscopy Database, formerly the ASGE database) has a list of commonly-used terms that it recognizes when a unique combination of initial letters is typed and thus reduces typing somewhat. However, the lists are long and not visible to the physician when he begins to fill in the blanks. These systems are easy to program, but difficult for physicians to use

0195-4210/91/$5.00

because few are proficient at typing. As a result, physicians often fill in paper forms, and use a secretary to transfer data from paper to a computer. Copying data is a major source of errors and requires a skilled (hence expensive) transcriptionist. Prose reports produced by these systems seem stilted. Since they offer no advantage over dictation, these systems are not widely accepted.

Menu-driven endoscopic data entry systems (such as the Pentax Reporter) are more difficult to program, but can be used by both typists and non-typists. Typists can enter a series of initial letters that uniquely specifies their choice. Non-typists can use arrow keys or a mouse to move highlighting along the list to choose the desired alternative. Because these systems are more efficient than dictation they've achieved some success, but many physicians are still hesitant to explore such methods of data entry. Also, the prose they produce is stilted.

Voice-driven endoscopic reporting software is the most difficult to program, but has the potential to solve the above problems while providing the physician with a means of data entry with which he/she is comfortable. In such a system (Kurzweil VoiceMED(tm) Systems), the physician works at a computer screen displaying a report text with blanks imbedded in it, as well as a list of words to fit in each of the blanks, and instructions to speak the words into a microphone. The computer analyzes the speaker's utterances and picks the words whose sounds are most similar. The physician's choices are then placed in the blank. The physician also uses voice to edit the report form on the screen, to print a report, and to capture images from an endoscope. Prose reports produced by such systems approach natural language because a complex structure can be associated with each word.

Computer systems for voice recognition are currently available, and such systems have been developed for radiology and other specialties (KurzweilVoiceRAD;...VoiceEM;...VoicePATH(tm)) but software for endoscopy is currently under development. The development of such software acceptable to practicing physicians will require the integration of computer technology with endoscopic systems. The minimum requirements for such a system are 1) data entry, storage, association with images, and retrieval with voice or from menus; 2) speed; 3) compatibility with generally available computers using major databases and operating systems so that the computer and database systems are available for other tasks (data analysis, word processing, etc.); 4) provision for multiple users to enter, retrieve, edit, and store data at once (since most GI labs have several rooms); 5) portable systems; and 6) unstilted prose reports. The software would be in a processor-independent language such as C. VoiceGI satisfies all these criteria.

The potential advantages of computer-generated reports are only available if the system is used by physicians. A system will not be used unless it is more efficient than dictation in the use of a physician's time, no matter what the theoretical advantages. The advantages of this voice-driven system to the physician include familiarity with the method of data entry, immediate feedback of a legible prose report, and completion of documentation activities at a single sitting. There is no need to touch a computer. Indeed, the report could be entered by voice during the procedure, resulting in further savings of time.

Available methods of endoscopic data entry have been developed primarily by programers in consultation with endoscopists. The endoscopists' involvement was secondary both programmatically and personally. The system to be demonstrated was developed by Kurzweil and Dr. Cass, a practicing endoscopist with extensive programming knowledge and experience and tested at nine academic and private gastroenterology practices nationwide. This has vastly improved the development of this software and thus will increase the chance for physician acceptance of a voice-driven data management system.

3. Principal Programs and Items to be Demonstrated

1) Kurzweil VoiceGI Program: 386-based PC with 387 math coprocessor, 16MB high-speed 32-bit bus RAM, 40MB hard disk, 1.2MB floppy disk, acoustic/phonetic analyzer card, VGA card and monitor, and TargaPlus 64 board; 2) Kurzweil Knowledgebase Editor Program; 3) Pentax Electronic Videoendoscope 2900; and 4) Pentax Electronic Videoprocessor EPM-3000.