Military Applications of Natural Language Processing and Software

James A. Rodger
Indiana University of Pennsylvania, USA

Tamara V. Trank
Naval Health Research Center, USA

Parag C. Pendharkar
Pennsylvania State University at Harrisburg, USA

EXECUTIVE SUMMARY

A preliminary feasibility study aboard U.S. Navy ships utilized voice interactive technology to improve medical readiness. A focus group was surveyed about reporting methods in health and environmental surveillance inspections to develop criteria for designing a lightweight, wearable computing device with voice interactive capability. The voice interactive computing device included automated user prompts, enhanced data analysis, presentation and dissemination tools in support of preventive medicine. The device was capable of storing, processing and forwarding data to a server. The prototype enabled quick, efficient and accurate environmental surveillance. In addition to reducing the time needed to complete inspections, the device supported local reporting requirements and enhanced command-level intelligence. Where possible, existing technologies were utilized in creating the device. Limitations in current voice recognition technologies created challenges for training and user interface.

BACKGROUND

Coupling computer recognition of the human voice with a natural language processing system makes speech recognition by computers possible. By allowing data and commands to be entered into a computer without the need for typing, computer understanding of naturally spoken languages frees human hands for other tasks. Speech recognition by computers can also increase the rate of data entry, improve spelling accuracy, permit remote access to databases utilizing wireless technology and ease access to computer systems by those who lack typing skills.

Variation of Speech-to-Text Engines

Since 1987, the National Institute of Standards and Technology (NIST) has provided standards to evaluate new voice interactive technologies (Pallett, Garofolo & Fiscus, 2000). In a 1998 broadcast

news test, NIST provided participants with a test set consisting of two 1.5-hr subsets obtained from the Linguistic Data Consortium. The task associated with this material was to implement automatic speech recognition technology by determining the lowest word error rate (Herb & Schmidt, 1994; Fiscus, 1997; Greenberg, Chang & Hollenback, 2000; Pallett, 1999). Excellent performance was achieved at several sites, both domestic and abroad (Przybocki, 1999). For example, IBM-developed systems achieved the lowest overall word error rate of 13.5%. The application of statistical significance tests indicated that the differences in performance between systems designed by IBM, the French National Laboratories' Laboratoire d'Informatique pour la Mechanique et les Sciences de l'Ingenieur and Cambridge University's Hidden Markov Model Toolkit software were not significant (Pallett, Garfolo and Fiscus, 2000). Lai (2000) also reported that no significant differences existed in the comprehension of synthetic speech among five different speech-to-text engines used. Finally, speaker segmentation has been used to locate all boundaries between speakers in the audio signal. It enables speaker normalization and adaptation techniques to be used effectively to integrate speech recognition (Bikel, Miller, Schwartz, & Weischedel, 1997).

Speech Recognition Applications

The seamless integration of voice recognition technologies creates a human-machine interface that has been applied to consumer electronics, Internet appliances, telephones, automobiles, interactive toys, and industrial, medical, and home electronics and appliances (Soule, 2000). Applications of speech recognition technology are also being developed to improve access to higher education for persons with disabilities (Leitch & Bain, 2000). Although speech recognition systems have existed for two decades, widespread use of this technology is a recent phenomenon. As improvements have been made in accuracy, speed, portability, and operation in high-noise environments, the development of speech recognition applications by the private sector, federal agencies, and armed services has increased.

Some of the most successful applications have been telephone based. Continuous speech recognition has been used to improve customer satisfaction and the quality of service on telephone systems (Charry, Pimentel & Camargo, 2000; Goodliffe, 2000; Rolandi, 2000). Name-based dialing has become more ubiquitous, with phone control answer, hang-up, and call management (Gaddy, 2000a). These applications use intuitive human communication techniques to interact with electronic devices and systems (Shepard, 2000). BTexact Technologies, the Advanced Communications Technology Centre for British Telecommunications, (Adastral Park, Suffolk, England) uses the technology to provide automated directory assistance for 700 million calls each year at its UK bureau (Gorham & Graham, 2000). Studies in such call centers have utilized live customer trials to demonstrate the technical realization of full speech automation of directory inquiries (McCarty, 2000; Miller, 2000). Service performance, a study of customer behavior and an analysis of service following call-back interviews suggest user satisfaction with the application of speech automation to this industry (Gorham & Graham, 2000).

Speech recognition technologies could expand e-commerce into v-commerce with the refinement of mobile interactive voice technologies (McGlashan, 2000; Gaddy, 2000b; Pearce, 2000). As an enabler of talking characters in the digital world, speech recognition promises many opportunities for rich media applications and communications with the Internet (Zapata, 2000). Amid growing interest in voice access to the Internet, a new Voice-extensible Markup Language (VoiceXML™, VoiceXML Forum) has surfaced as an interface for providing Web hosting services (Karam & Ramming, 2000). VoiceXML promises to speed the development and expand the markets of Web-based, speech recognition/synthesis services as well as spawning a new industry of "voice hosting." This model will allow developers to build new telephone-based services rapidly (Thompson & Hibel, 2000). The voice-hosting service provider will lease telephone lines to the client and voice-enable a specific URL, programmed in VoiceXML by the client. This model will make it possible to build speech and telephony services for a fraction of the time and cost of traditional methods (Larson, 2000).

Related Content

Assessment of End-User Computing from an Organizational Perspective

François Bergeron, Suzanne Rivardand Louis Raymond (1993). *Information Resources Management Journal (pp. 14-25).*

www.irma-international.org/article/assessment-end-user-computing-organizational/50972

A Process Approach for Selecting ERP Software: The Case of Omega Airlines

Jacques Verville (2003). *Annals of Cases on Information Technology: Volume 5 (pp. 26-44).*

www.irma-international.org/article/process-approach-selecting-erp-software/44531

RFID: New Technology on the Horizon for IT Majors

Eric Puffenbarger, Faye P. Teerand S. E. Kruck (2008). *Information Communication Technologies: Concepts, Methodologies, Tools, and Applications (pp. 3730-3743).*

www.irma-international.org/chapter/rfid-new-technology-horizon-majors/22911

Organizational Readiness for Successful Knowledge Sharing: Challenges for Public Sector Managers

W.A. Taylorand G.H. Wright (2004). *Information Resources Management Journal (pp. 22-37).*

www.irma-international.org/article/organizational-readiness-successful-knowledge-sharing/1254

Customer Relationship Management and Knowledge Discovery in Database

Jounghae Bang, Nikhilesh Dholakiam, Lutz Hameland Seung-Kyoon Shin (2009). *Encyclopedia of Information Science and Technology, Second Edition (pp. 902-907).*

www.irma-international.org/chapter/customer-relationship-management-knowledge-discovery/13682