S.K. Andersen et al. (Eds.)
IOS Press, 2008

Voice-controlled Data Entry in Dental Electronic Health Record

Miroslav NAGY[1,a], Petr HANZLICEK[a], Jana ZVAROVA[a], Tatjana DOSTALOVA[b], Michaela SEYDLOVA[b], Radim HIPPMAN[b], Lubos SMIDL[c], Jan TRMAL[c] and Josef PSUTKA[c]

[a] *EuroMISE Center, Department of Medical Informatics, Institute of Computer Science AS CR, v.v.i., Prague, Czech Republic*

[b] *Department of Paediatric Stomatology, 2nd Medical School, Charles University in Prague*

[c] *Department of Cybernetics, University of West Bohemia in Pilsen*

Abstract. The EuroMISE Center focuses on new approaches in the field of electronic health record (EHR). Among others, the structured health documentation in dentistry in the form of an EHR is being systematically studied. This paper describes the evolution of the EHR developed at the EuroMISE Center named MUDRLite and its graphical component for dentists called DentCross. The summary of features of the DentCross component is followed by a brief description of automatic speech recognition and an ASR module. The problems with data insertion into EHR during examination of a dental patient lead to further research in the area of the automatic speech recognition in medical practice. Cooperation of engineers, informaticians and dental physicians resulted in an application called DentVoice which is a successful application of the ASR module and the DentCross component of the MUDRLite EHR. The junction of voice control and graphical representation of dental arch makes hand-busy activities in dental praxis easier, quicker and more comfortable. This will result in a better quality of the data stored in a structured form in dental EHR, thus enabling better decision making and use of decision support systems.

Keywords. automatic speech recognition, electronic health record, dental medicine

Introduction

Patients' health data are an integral part of dental healthcare documentation and they are constantly becoming more important. Besides insurance companies, information about patient's health status is required in practice in order to improve health care quality and decision making. There are still some obstacles that have to be surpassed.

One of them is the lack of rich structured data in dental medicine. Information still appears in the form of free-text-based documentation. Data structuring is on a relatively low level, mostly limited to tables, containing more or less standardized symbols (e.g.

[1] Corresponding Author: Miroslav Nagy, EuroMISE Center, ICS AS CR v.v.i., Pod Vodarenskou vezi 2, 182 07 Prague 8, Czech Republic; E-mail: nagy@euromise.cz

"/" for caries, "-" for pulpitis, or "x" for a tooth to be extracted), placed in the section corresponding to a particular tooth. The main disadvantage of using such a system is the loss of detailed information on the localization, the size and the character of hard dental tissues defects. Further information on an oral cavity is not sufficiently structured as well. Information concerning changes of oral cavity mucosa, periodontitis, orthodontic anomalies, preventive oncological examinations, etc. is described in a limited space of one line of the form or together as other findings in the form of a free text.

It has been a common situation in dental practice that during check-up the dentist dictated patient's data to a nurse, who filled out paper based forms. Nowadays, computerization in dental practice brought major changes in data entry. Nurse enters patient's data into a specialized computer application. As an example of a dental application used in Czech Republic we can mention the PC Dent software by DialogMIS [1]. The PC Dent is a medical information system for dentists, containing textual patients' records, support of scheduling and a graphical tool for dental status check-up and treatments. The PC Dent is one of the most elaborated medical software systems in dentistry available on the Czech market and is received quite well by its users.

Entering data into a computer during the examination of a patient requires either costly changing of sterile gloves, which is also time-consuming, or the participation of a nurse, entering the dictated data. These problems could be possibly overcome by use of the automatic speech recognition in EHR. Voice commands usage has been examined since 1990's [2], [3] as a convenient replacement of computer keyboard and mouse. Necessity of using human voice to control a computer or other device arose in typical hands-busy environments such as surgery [4] or dentistry [5].

1 DentCross Component and Automatic Speech Recognition in MUDRLite EHR for Dentistry

1.1 MUDR EHR and Dental Knowledge Base

Several European projects, international standards and long-lasting cooperation with physicians helped us to develop a pilot EHR system called MUDR (MUltimedia Distributed Record) [6], [7]. It is based on the three-tier architecture with an unusual data-storing approach based on two main structures described by tools of a graph theory, named knowledge base and data-files.

In the years 2004 and 2005 the dental knowledge base was created. Dentists and informaticians from the joint workplace of the EuroMISE Center produced a model in the form of a knowledge tree as a part of the knowledge base of the MUDR EHR. The knowledge tree comprised of basic information about patient, family history, social history, personal history, information about medication and exhaustive information about patient's oral cavity status form the dentist's point of view. Approximately 1000 dental concepts were structured during this process. The knowledge tree served as a model basis for further dental EHR created at the EuroMISE centre.

1.2 MUDRLite

The MUDRLite is an EHR inspired by the MUDR EHR. Its architecture is based on two layers. The first one is a relational database. Currently, MS SQL server versions 7 and 2000 are supported. The second layer is a MUDRLite User Interface (MUDRLite UI) running on a Windows-based operating system using Microsoft .NET Framework.

The database schema corresponds to the particular needs and varies therefore in different environments, as opposed to the fixed database schema in the MUDR data layer. MUDRLite universality is based on a different approach. The core of MUDRLite – MUDRLite Interpreter – is able to handle various database schemas. This feature often simplifies the way of importing old data stored with other databases or files.

All the visual aspects and the behaviour of the MUDRLite UI are completely described by an XML configuration file. The end-user can see a set of forms with various controls placed on them by appropriate XML elements. MUDRLite operates as a kind of commands' interpreter; it processes the instructions encoded in the so called MLL Language as described in [8] and manipulates the database layer as well as the visual aspects of the MUDRLite UI.

1.3 DentCross Component

To gain MUDRLite's user-acceptance in the field of dental medicine we have developed an advanced component representing the dental cross, which is a crucial part of healthcare documentation in dental medicine. Its development was motivated by clinical practice in the dental medicine domain. The DentCross component was included dynamically into the user interface in a form of user-defined custom component.

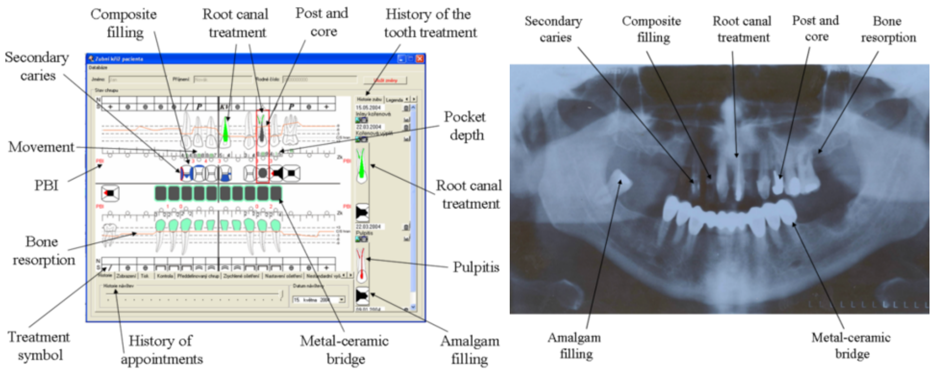

Figure 1. The DentCross Component – a Case Study Example and the Corresponding Tomograph Picture.

The DentCross component is implemented as a stand-alone library DentCross.dll, completely developed for the .NET Framework platform using the Microsoft Visual Studio .NET 2003 development tool.

A user-defined component is inserted by the custom element of the MLL Language with the following mandatory attributes: "dll" specifying the name of the assembly the component is implemented in, and "class" specifying the name of the main class of the included component.

For the end user, the DentCross component looks like a kind of a dental panoramic tomograph. This component is fully interactive and enables to record fully structured dental medicine information that can be inserted user-friendly by mouse or keyboard. A dentist can choose among about 60 different actions [9], treatment procedures or tooth parameters that are displayed graphically and lucidly.

The Figure 1 depicts a screenshot of the DentCross component showing dental status of a case study patient and X-ray image representing the same situation. Apparently, the graphical representation of the dental arch produced by DentCross component is as illustrative as the X-ray image and acceptable by the patient. Modern dental chairs have an integrated monitor that can be viewed by both the dentist and the patient, thus improving mutual trust and as a result lessening patient's stress.

1.4 Automatic Speech Recognition Module

The automatic speech recognition (ASR) engine was implemented by the Department of Cybernetics, University of West Bohemia in Pilsen [10]. The module was implemented as a standalone application, running in server-mode in background.

The server communication protocol is proprietary, running on the top of TCP/IP stack. The communication protocol enables the startup and shutdown of the recognition process, the run-time configuration of the recognition task and also receiving the recognized phrases by a client.

The ASR system is speaker-independent and is based on a statistical approach. The goal was to implement a recognition module featuring high accuracy yet able of very fast operation. It comprises a front-end, an acoustic model, and a decoding block as can be seen on Figure 2.

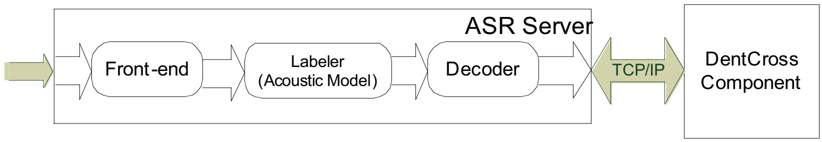

Figure 2. Automatic Speech Recognition functionality.

1.4.1 Front-end

The speech signal is digitized at 8 kHz sample rate and converted to the u-law format. Then the pre-emphasized acoustic waveform is segmented into 25 milliseconds frames with 15 ms overlap. A Hamming window is applied to each frame and static PLP cepstral coefficients (PLP_CCs) are computed. Moreover, the delta (first order derivatives) and the delta-delta (second order derivatives) coefficients are computed and appended to the feature vector.

1.4.2 Acoustic Model

The acoustic model is based on modeling of triphones. Each individual triphone is represented by a three state left-right HMM (Hidden Markov Model) with a continuous output probability density function assigned to each state. Each density function is approximated as a mixture of multivariate Gaussians with diagonal covariance matrices.

The number of mixture components for each state was obtained experimentally. Since a variety of noise sounds (e.g. loud breath, click on the microphone and noise of an audio channel) can appear in every utterance, a set of noise HMM units was developed and trained.

1.4.3 Decoder Block

The decoder uses a Viterbi search technique together with efficient beam pruning [11]. The search is performed on a crossword context dependent HMM state network. The state network is generated from EBNF (extended Backus–Naur form) metasyntax notation described in JSGF (Java Speech Grammar Format) format. The whole recognition network is constructed from one or more together connected grammars.

To improve the robustness, the decoder contains a rejection technique to refuse incoming utterances that are not accepted by the grammar (or grammars). In this case the ASR engine informs the client that the utterance was not successfully recognized.

2 DentVoice Application

Interconnection of the ASR module and the DentCross component of the MUDRLite EHR resulted in an application called DentVoice. Junction of voice control and graphical representation of dental arch makes hand-busy activities in dental praxis easier, quicker and more comfortable. Dentists were involved in the whole development process; therefore manipulation with the DentCross component was designed to be as easy as possible.

The prototype application consists of a DentCross component with integrated TCP/IP client of the ASR server and a voice commands definition file. The ASR client uses a DentCrossHandler class that implements all functionality of the DentCross component.

The speech recognition is activated immediately after DentCross component start-up. The recognition process can be paused or stopped by a special voice command or using the user interface. Voice commands can be divided in two groups: global manipulating commands and context dependent commands.

Global commands are always available and are designed to manipulate the recognition process (pause, resume and stop) and to close message boxes opened by the application to warn the user.

Context dependent commands rely on the current state of the DentCross component and can be further divided into 33 command groups because it can fall into one of 33 states (e.g. tooth treatment, caries placement, caries type, root canal treatment material).

3 Discussion

First experience with the DentVoice application showed that the need of abbreviated form of commands is inevitable. Since the main goal of the voice controlled approach is to make the data entry easier and faster, further shortcuts should be found. On the other side, some minimal length of commands should be kept, as the length of commands (the longer the better) plays significant role in improving the recognition hit rate. Abbreviated forms of positions on the tooth such as D for distal or B for buccal, are very similar from the phonetic point of view and therefore easily mistakable by the ASR module. Further improvement of the accuracy of the ASR module has been achieved by having voice commands separated in state dependent groups, thus limiting the number of possible commands in a particular state.

Forthcoming development regards usage of computer synthesized speech. The DentCross will read the actual status of the patient's teeth stored in the database and the dentist will just check if the information corresponds with reality. If so, the check-up ends without any necessary action, otherwise, the reading process is paused and the dentist dictates the divergence he or she found. In order to continue the check-up the reading process is resumed. The synthesized computer voice will serve as a feedback to the dentist that can stay focused and look in patient's mouth only to validate the correctness of the stored data.

DentCross component has been successfully used in forensic dentistry [12] to improve identification of disaster victims. Voice controlled version of the DentCross, DentVoice application, will be getting further improvement in data entry during the dead corpses inspection.

Acknowledgement

This paper was supported by the project 1ET200300413 of the Academy of Sciences of the Czech Republic.

References

[1] Dialog MIS, http://www.dialogmis.cz/english/index.html (last accessed November 6, 2007).
[2] Grasso MA. Automated speech recognition in medical applications. MD Computing 1995; 12(1):16-23.
[3] Grasso MA, Ebert D, Finin T. Acceptance of a speech interface for biomedical data collection. Proc AMIA Annu Fall Symp 1997; 739-43.
[4] Reichenspurner H, Damiano RJ, Mack M, Boehm DH, Gulbins H, Detter Ch, Meiser B, Ellgass R, Reichart B. Use of the voice-controlled and computer-assisted surgical system ZEUS for endoscopic coronary artery bypass grafting. J Thorac Cardiovasc Surg, 1999; 118: 11-16.
[5] Drevenstedt GL, McDonald JC, Drevenstedt LW. The role of voice-activated technology in today's dental practice. J Am Dent Assoc, 2005; 136(2): 157-161.
[6] Hanzlicek P, Spidlen J, Nagy M. Universal electronic health record MUDR. Stud Health Technol Inform 2004; 105: 190-201.
[7] Hanzlicek P, Spidlen J, Heroutova H, Nagy M. User interface of MUDR electronic health record. Int J Med Inform, 2005; 74(2-4): 221-7.
[8] Spidlen J, Hanzlicek P, Zvarova J. MUDRLite: Health record tailored to your particular needs. Stud Health Technol Inform 2004; 105: 202-9.
[9] Spidlen J, Pies M, Teuberova Z, Nagy M, Hanzlicek P, Zvarova J, Dostalova T. MUDRLite – an electronic health record applied to dentistry by the usage of a dental-cross component. IFMBE Proceedings Vol.11; 2005.
[10] Muller L, Psutka J, Smidl L. Design of speech recognition engine. Lect Notes Comput Sci 2000; 1902: 259-64.
[11] Psutka J, Muller L, Matousek J, Radova V. Mluvíme s počítačem česky. Praha, Academia, 2006.
[12] Zvarova J, Dostalova T, Hanzlicek P, Teuberova Z, Nagy M, Pies M, Seydlova M, Eliasova H, Simkova H. Electronic health record for forensic dentistry. Methods Inf Med 2008; 47: 8-13.