Early Diagnosis of Alzheimer's Type Dementia Using Continuous Speech Recognition

Vassilis Baldas[1], Charalampos Lampiris[2], Christos Capsalis[2], and Dimitrios Koutsouris[1]

[1] Biomedical Engineering Laboratory, National Technical University of Athens, Greece
[2] Wireless Communication Lab, National Technical University of Athens, Greece
baldas@biomed.ntua.gr, clamp@mail.ntua.gr, ccaps@central.ntua.gr, dkoutsou@biomed.ntua.gr

Abstract. One of the most important social problems that many of the developed countries face is the constant rise of the percentage of the elderly population values. A major health issue affecting this part of the population is the appearance of dementia of the Alzheimers type (AD) which is the most common case of dementia, affecting around 50 million people worldwide. In addition, the frequency of the AD related cases is expected to grow three times over the next 50 years.

The proposed AD status monitoring system (ADSMS) is processing a person's speech habits to train itself and extracts specific statistic parameters. After that necessary training process, it constantly monitors and analyzes new spontaneous speech data, in order to classify them. In this way, it is possible for the ADSMS to predict possible signs of AD that need further investigation with the common methods of diagnosis by a physician.

Keywords: Dementia, Alzheimer, Continuous Speech Recognition, Speech Analysis.

1 Introduction

One of the most important social problems that many of the developed countries face, is population aging. Demographics show that there is a constant rise in the percentage of the elderly population values. One of the main health issue affecting this part of the population is the appearance of dementia [1]. The dementia of Alzheimer s type (AD) is the most common case of dementia and affects around 50 million people worldwide and the frequency of the AD related cases is expected to grow three times over the next 50 years [2].

Alzheimer's and other dementias are also a great social and fiscal burden for the nations, therefore an early detection will be of significant importance for both the patients and the health and social care system. Early diagnosis leads to early access to pharmacological treatment, access to information and education

J. Lin and K.S. Nikita (Eds.): MobiHealth 2010, LNICST 55, pp. 105–110, 2011.

for everyone involved, organization of the counseling and community support, cognitive training and even lifestyle advice. The benefits of early institutionalization and improved physical and mental health will also result in better public health and potentially lower costs [3].

Some work has been done on the field of early detection of AD using electroencephalographic (EEG) recordings and classifiers [4], and also by analyzing the MRI volumes in search of volumetric atrophy of the gray matter (GM) in areas of neocortex of AD patients. Most of these techniques use machine learning methods to classify brain images (MRI or fMRI) in order to short and discriminate the characteristics of normal vs neuropathological subjects [5]. In general, most of the previous work on early detection of AD is based on analysis of medical signals and images, and on gene information extraction [6].

Some researchers though, managed to extract information about a person's cognitive status from analyzing his written or spoken language [7] [8]. The described solution uses these methods as the basic theory behind the AD detection system.

1.1 The Proposed Method for Early Detection of AD

Nowadays, personal computers along with various multimedia and communication systems are present in every home. This fact can be used to design a pervasive long-term AD status monitoring system, based on continuous speech recognition and analysis. The person whose mental and cognitive level we are interested on tracking, is hereinafter called *subject*.

The AD status monitoring system (ADSMS) extracts specific statistical measurements using data from the subject's speech analysis, to train itself. After the training process is complete it constantly analyzes new data that are entered to the system in order to classify them. Thus, it is possible for the ADSMS to predict possible signs of AD that need further investigation with the common methods of diagnosis by the subject's physician.

2 Description of the Proposed System

The overall architecture of the proposed system and the main modules that contribute to the final result, are shown on Fig 1.

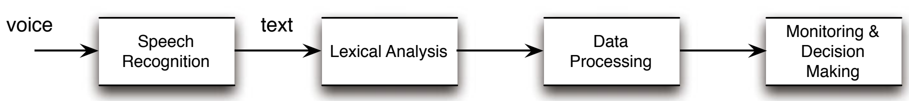

Fig. 1. The basic functions of the proposed system

ADSMS consists of three primary modules and each of them is composed of further subsystems. These parts are presented below.

2.1 Speech Recognition

The first part of ADSMS is the speech recognition module. This is the subsystem that is responsible for collecting the voice samples from the subject via a microphone or a network of multiple microphones, and transforming them to raw text that is later going to be further analyzed. Many methods in bibliography have been presented on accurate and fast speech recognition [9]. Two of the most used technologies on the field are hidden Markov modeling (HMM) and dynamic programming search techniques for large-scale networks [10]. Some of these works manage to succeed in translating speech to text even on noisy environments, and even with mobile devices as the processing platform.

For ADSMS, we propose a voice recognition system like the one described on [9], that uses multiple microphones in a room to collect the samples of the subject's voice and then convert them to text. The speech recognition subsystem of ADSMS is considered to be the most trivial part, as the research that has been done on this area is impressive and the results are satisfying.

2.2 Lexical Analysis

We begin by defining the concept of *Speech Matrix*, that will be used further on. Assuming we have a total of n words in the speech recognizer (SR) database, we map each of those words to a respective place in a speech vector of size $1 \times n$. Assuming also that the SR runs constantly each day, at the end of each day the respective speech vector for the i-th day will be of this form: $W_i = \big(t_{i,1}\ t_{i,2} \cdots t_{i,n}\big)$ where $t_{i,j}$ is the number of times the j-th word is spoken on the i-th day.

For k days, we can represent the total speech body of the subject by the speech matrix:

$$W_k = \begin{pmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,n} \\ \vdots & & & \vdots \\ t_{k,1} & t_{k,2} & \cdots & t_{k,n} \end{pmatrix}$$

At the same time, all the sentences are analyzed for *part-of-speech-tagging* (POST) that is required on later stages for the analysis. Words are classified based on eight parts of speech: verb (V), noun (N), pronoun (PN), adjective (AJ), adverb (ADV), preposition (PP), conjunction (CJ), and interjection (IJ). Using the algorithm described on [11] we compute for the i-th day a POST vector $POST_i = \big(V_i\ N_i\ PN_i\ AJ_i\ ADV_i\ PP_i\ CJ_i\ IJ_i\big)$ where χ_i is the number of words that were identified as belonging to the χ part-of-speech. Following the same rules as with the speech matrix, we can build a POST matrix for k days ($POST_k$).

Using the data from the W_k and the $POST_k$ matrices, we can extract important statistic values that provide information about the subject's cognitive status.

2.3 Interpretation of Results

Some previous work has been done on correlating spontaneous speech statistical patterns with AD [7] [8]. The authors use lexical analysis of the patient's words to get results of the severity of the AD. The stylometric measures that can be used to analyze the subject's W_k and $POST_k$ matrices are borrowed from the linguistics, and the ones that will be used by ADSMS are shown on Table 1.

Table 1. The stylometric measures that can be used to distinguish AD patients from healthy subjects

Symbol	Definition
Total Words (N)	Number of words spoken
Vocabulary Size(Voc)	Number of different words
Noun rate (N_{rate})	$\frac{nouns}{N}$
Pronoun rate (P_{rate})	$\frac{pronouns}{N}$
Adjective rate (A_{rate})	$\frac{adjectives}{N}$
Verb rate (V_{rate})	$\frac{verbs}{N}$
Type-Token Ratio (TTR)	$\frac{Voc}{N}$
Brunét's index (W)	$N^{Voc^{-0.615}}$
Single Vocabulary (Voc^{single})	Number of different words spoken once
Honorés Statistic (R)	$\frac{100 \log N}{1-\frac{Voc^{single}}{Voc}}$

Bucks et al. [8], found that AD patient's have different speech patterns than normal subjects. These findings, along with findings from [7] are summed up on Table 2.

Each of the metrics can be easily computed from the subject's W_k and $POST_k$ matrices by manipulating and combining the rows.

2.4 Training of the ADSMS for Each Subject and Implementation Notes

Different subjects present different metric values, so it is important to train the ADSMS per person in order to obtain the best results. The training process must be done while the subject is in good cognitive condition so that the system is able to have the normal condition's results, which are the control values. Constant periodical comparison between these normal values and the new data will make it possible to monitor the subject's cognitive status and provide possible alerts for further investigation by his physician.

Table 2. Stylometric results. AD vs. healthy groups.

Metric	Result
Noun rate (N_{rate})	$N_{rate}^{AD} \leq N_{rate}^{normal}$
Pronoun rate (P_{rate})	$P_{rate}^{AD} \geq P_{rate}^{normal}$
Adjective rate (A_{rate})	$A_{rate}^{AD} \geq A_{rate}^{normal}$
Verb rate (V_{rate})	$V_{rate}^{AD} \geq V_{rate}^{normal}$
Type-Token Ratio (TTR)	$TTR^{AD} \leq TTR^{normal}$
Brunét's index (W)	$W^{AD} \leq W^{normal}$
Honorés Statistic (R)	$R^{AD} \leq R^{normal}$

Aging of the subjects is also expected to bring changes to some metrics [7] [8], so it is clear that the control (normal) values have to be updated when this change affects the results.

The speech recognition and data collecting module of the ADSMS has to work constantly in order to capture the data. The heavy work of data processing, analysis, monitoring and decision making can be done on a personal computer during the idle time as a background process.

3 Conclusions and Future Work

We have presented a method for constant monitoring of a subject's speech that can analyze the lexical data and decide on whether or not his cognitive status may be deteriorating due to AD. The described system can be installed on every space, and it's technological requirements are trivial.

Future additions and improvements may include the integration of the whole system into one wearable device that is constantly on the subject's body, and the extension of this research into constant monitoring of other health conditions to investigate if they present similar patterns.

References

1. De Castro, A.K.A., Pinheiro, P.R., Pinheiro, M.C.D.: Applying a decision making model in the early diagnosis of alzheimer's disease. In: Yao, J., Lingras, P., Wu, W.-Z., Szczuka, M.S., Cercone, N.J., Ślęzak, D. (eds.) RSKT 2007. LNCS (LNAI), vol. 4481, pp. 149–156. Springer, Heidelberg (2007)
2. Ramrez, J., Grriz, J., Salas-Gonzalez, D., Romero, A., Lpez, M. lvarez, I., Gmez-Ro, M.: Computer-aided diagnosis of alzheimer's type dementia combining support vector machines and discriminant set of features. In: Information Sciences (2009)
3. Todd, S., Passmore, P.: Alzheimers disease, the importance of early detection. European Neurological Review (2009)

4. Cichocki, A., Shishkin, S.L., Musha, T., Leonowicz, Z., Asada, T., Kurachi, T.: Eeg filtering based on blind source separation (bss) for early detection of alzheimer's disease (ENY-ARTICLE-2009-243. AZ I07/2005/I-022) (Sie 2004). 9p Clinical Neurophysiology 116(3), 729–737 (2005)
5. Savio, A., García-Sebastián, M., Hernández, C., Graña, M., Villanúa, J.: Classification results of artificial neural networks for alzheimer's disease detection. In: Corchado, E., Yin, H. (eds.) IDEAL 2009. LNCS, vol. 5788, pp. 641–648. Springer, Heidelberg (2009)
6. Schenk, D., Barbour, R., Dunn, W., et al.: Immunization with amyloid-beta attenuates alzheimer-disease-like pathology in the pdapp mouse. Nature 400, 173–177 (1999)
7. Thomas, C., Cercone, N.: Automatic detection and rating of dementia of alzheimer type through lexical analysis of spontaneous speech. In: Proc. of IEEE ICMA (2005)
8. Bucks, R., Singh, S., Cuerden, J.M., Wilcock, G.K.: Analysis of spontaneous, conversational speech in dementia of alzheimer type: Evaluation of an objective technique for analysing lexical performance (2000)
9. Yamamoto, K., Masatoshi, T., Nakagawa, S.: Privacy protection for speech signals. Procedia - Social and Behavioral Sciences 2(1), 153–160 (2010); The 1st International Conference on Security Camera Network, Privacy Protection and Community Safety (2009)
10. Jiang, H.: Confidence measures for speech recognition: A survey. Speech Communication 45(4), 455–470 (2005)
11. Schmid, H.: Probabilistic part-of-speech tagging using decision trees. In: Proceedings of the International Conference on New Methods in Language Processing, Manchester, UK (1994)