35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Early Detection of Cognitive Impairment in the Elderly Based on Bayesian Mining Using Speech Prosody and Cerebral Blood Flow Activation

Shohei Kato, *Member, IEEE*, Hidetoshi Endo, Akira Homma, Takuto Sakuma, and Keita Watanabe

Abstract— **With the aim of providing computer aided diagnosis of dementia, we have developed a non-invasive screening system of the elderly with cognitive impairment. In our previous research, we have studied two data-mining approaches by focusing on speech-prosody and cerebral blood flow (CBF) activation during cognitive tests. On the power of these research results, this paper presents a prosody-CBF hybrid screening system of the elderly with cognitive impairment based on a Bayesian approach. The system is constructed by SPCIR (Speech Prosody-Based Cognitive Impairment Rating) based cutoff as the 1st screening, and, as the 2nd screening, two-phase Bayesian classifier for discriminating among elderly individuals with three clinical groups: elderly individuals with normal cognitive abilities (NC), patients with mild cognitive impairment (MCI), and Alzheimer's disease (AD). This paper also reports the screening examination and discusses the cost-effectiveness and the discrimination performance of the proposed system for early detection of cognitive impairment in elderly subjects.**

I. Introduction

ABRUPT increase in elderly patients with dementia due to growing super-aging society comes to public attention in developed countries. Research and development of new dementia medications are accelerated. Development of the early detection methods for dementia that are both sensitive and specific is also very important as a diagnostic tool.

To screen for dementia and cognitive impairment, a questionnaire test such as Mini-Mental State Examination (MMSE) [1], Revised Hasegawa's Dementia Scale (HDS-R) [2], Clinical Dementia Rating (CDR) [3], and Memory Impairment Screen (MIS) [4], is commonly used in addition to a neurophysiological test (e.g., using MRI [5], FDG-PET [6], and CSF biomarkers [7]). Questionnaire tests have some disadvantages and their use is limited in the clinic. The MMSE, HDS-R, and CDR are more time-consuming than a general practitioner's consultation. In general, the questionnaire cannot completely dismiss the influence of education, social class, and gender difference on the results. In addition, there is a possibility that practitioner subjectivity may affect the scoring. Thus, we believe that the development of a simple, non-invasive examination that is objective and combined with a physiological test could enables the early detection of dementia in a broad population.

In our previous study, we have studied novel approaches for the early detection of cognitive impairment in the elderly, in which we focused on the prosodic features of speech sound during the subject's answers to the questionnaire; the first was to detect signal and prosodic signs of cognitive impairment [8], and the second was to take a measurement of cerebral blood flow (CBF) [9]. The first method had an advantage that enables everyone to check his/her own cognitive ability anywhere because of using speech signals only. The method is effective for the first step of screening for dementia, but, it has limitations of the reliability because the method does not measure brain function. On the other hand, a neurophysiological test, such as using MRI, FDG-PET, and CSF biomarkers, imposes severe constraint on a subject, for instance, pain at obtaining cerebral spinal fluid, radiation exposure, physical restraint and so on. This is a disadvantage in early screening, which should cover all elderlies. As the second step, we then focused on functional near-infrared spectroscopy (fNIRS) as a brain function measurement system, which can eliminate physical restraint from a subject by non-invasive procedures, and developed a prototype for computer-aided diagnosis of cognitive impairment in the elderly with the use of fNIRS signals during cognitive tests by learning a two-phase Bayesian classifier for discriminating among elderly individuals with three clinical groups.

In this paper, we present a two-phase screening system for cognitive impairment in the elderly by combination of dual signal processing techniques using speech prosody and fNIRS signals. In addition, we addressed the effectiveness of the proposed method in discriminating among elderly individuals with normal cognitive abilities (NC), patients with mild cognitive impairment (MCI), and Alzheimer's disease (AD).

Manuscript received February 7, 2013; revised April 16, 2013. This work was supported in part by the Ministry of Education, Culture, Sports, Science, and Technology in Japan, Grant–in–Aid for Scientific Research under grant #25280100 and #25540146, and part by the SENTAN, Japan Science and Technology Agency (JST), and part by the Adaptable and Seamless Technology Transfer Program through target-driven R&D, JST, and part by the Program to Promote Technology Transfer and Innovation through Collaboration between Universities, JST, and part by the Suzuken Memorial Foundation.

S. Kato, T. Sakuma, and K. Watanabe are with the Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466–8555, Japan (phone/fax: +81-52-735-5625; email: shohey@katolab.nitech.ac.jp; url: http://www.katolab.nitech.ac.jp).

H. Endo is with the Department of Comprehensive Geriatic Medicine, National Center for Geriatrics and Gerontology, 35 Gengo, Morioka-machi, Obu, Aichi 474-8511, Japan, (email: hendo@ncgg.go.jp).

A. Homma is with the Tokyo Dementia Care Research and Training Center, 1-12-1, Takaido Nishi, Suginami-ku, Tokyo 168-0071, Japan (email: ahomma96@dcnet.gr.jp).

TABLE I

A BREAKDOWN LIST OF PARTICIPANTS (N=48)

Age	64-70	71-75	76-80	81-85	86-92	Total
Male	3(2,0,1)	2(1,1,0)	4(3,1,0)	7(1,4,2)	2(0,0,2)	18(7,6,5)
Female	7(4,2,1)	6(4,2,0)	8(2,5,1)	5(2,1,2)	4(1,3,0)	30(13,13,4)
Subtotal	10(6,2,2)	9(6,3,0)	12(5,6,1)	13(3,5,5)	6(1,3,2)	48(20,19,9)

Value in bracket means the number of subjects in NC, MCI, AD clinical groups.

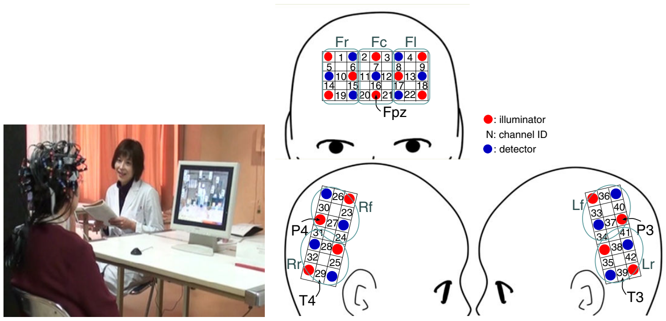

Fig. 1. Snapshot and channell arrangement of fNIRS measurement.

II. METHOD

A. Participants

Forty eight Japanese subjects (18 males and 30 females between the ages of 64 and 92 years) participated in this study. Table I shows the breakdown list of participants. In this study, all participants are clinically conditioned that CDR of a participant in MCI group and AD group corresponds to CDR0.5 and CDR1, respectively.

B. Cognitive Tests

To measure brain function of an elderly during various cognitive tests including HDS-R, we have made a block designed task, and then conducted simultaneous voice-fNIRS measurement during cognitive tests. Firstly a participant talks about the topics of hometown and childhood and answers for an HDS-R questionnaire test for ten minutes. And then, he/she does three reminiscence tasks (1. listening, 2. talking, 3. watching) and three working memory tasks (1. category recall, 2. reading span, 3. face recall) for twelve minutes. These six tasks are done for 60 seconds after rest gazing at a single point on the display for 60 seconds interval.

C. fNIRS Measurement

Functional near-infrared spectroscopy (fNIRS) can measure neural activity of the cerebral cortex using infrared rays that are safe to living organisms [10]. fNIRS monitors regional relative changes of oxy/deoxygenated hemoglobin concentration to measure cortical activation utilizing the tight coupling between neural activity and regional cerebral blood flow [11]. This measurement method requires only compact experimental systems and can eliminate physical restraint from a subject by non-invasive procedures (see Figure 1).

We used the fNIRS topography system FOIRE-3000 Near-Infrared Brain Function Imaging System (Shimadzu, Kyoto, Japan), which uses near-infrared light with wavelengths of 780, 805, and 830 nm. We set 16 illuminators and 15 detectors in lattice pattern to form 42 channels (CHs) (22 CHs on frontal lobe, 10 CHs on right parietal and temporal lobe, 10 CHs on left parietal and temporal lobe) shown in Figure 1.

TABLE II

CORRELATION BETWEEN SPCIR AND HDS-R

	$\text{SPCIR}_{\text{PCA-FSW-AIC}}$
# of regressors	42
Multiple-R^2	0.61
Adjusted-$\bar{R}^2$	0.54
S.E.	2.40
P-value	$< 2.2 \times 10^{-16}$

III. SIGNAL AND PROSODIC SIGNS OF COGNITIVE IMPAIRMENT IN THE ELDERLY SPEECH

The section describes speech prosody-based cognitive impairment rating, which is calculated by the multiple regression analysis of signal and prosodic signs in the elderly speech.

A. Extraction of Signal and Prosodic Signs

Speech has three components: prosody, tone, and phoneme. Past research indicates that the prosodic component has important non-verbal information such as emotional expressions [12]. In accordance with our hypothesis, cognitive impairment was observed in the elderly [13]. In this study, we considered 128 different acoustic correlates related to both segmental and suprasegmental information from speech signals. We used a computational data mining strategy based on a statistical-analytical approach. We extracted as many features as possible, and disregarded irrelevant features using a feature selection technique. These features were phrase-level statistics corresponding to

- fundamental frequency (F0) and their time-series behavior (13 features),
- formant and its time-series behavior (33 features),
- power envelope and its time-series behavior (22 features),
- lower twelve dimensions of mel-frequency cepstral coefficients (MFCCs) using 20 mel-scaled filters, and their time-series behavior (60 features).

Prosodic analysis was performed in 23-ms frames and passed through a Hamming window (1024 points). Voice waveforms (sampled at 44.1 kHz with 16 bits) were extracted using a short-time Fourier transform (STFT) every 11 ms.

B. SPCIR: Speech Prosody-Based Cognitive Impairment Rating

SPCIR adopts a novel approach of prosody-based speech sound analysis and a multivariate statistical technique. SP-

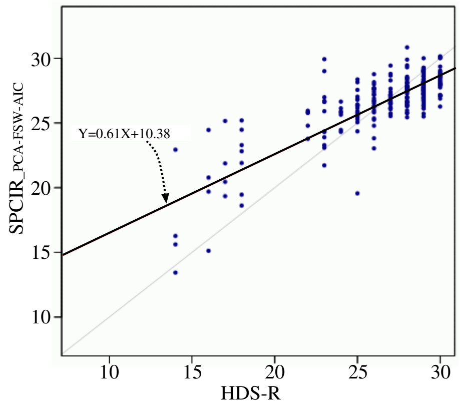

Fig. 2. Scatter plot of HDS-R and $\text{SPCIR}_{\text{PCA-FSW-AIC}}$.

CIR is calculated by multiple linear regression using prosodic features (as regressors) selected by the feature selection in forward stepwise method. We then analyzed the correlation between the HDS-R scores and SPCIRs obtained from speech prosody in elderly individuals using speech voice samples (N=48), each with 128 prosodic features.

Table II shows the results of the analysis and the scatter plots of HDS-R and the SPCIRs are shown in Figure 2. In this figure, $\text{SPCIR}_{\text{PCA-FSW-AIC}}$ denotes SPCIR with the PCA pre-processed forward stepwise method in combination with Akaike Information Criterion (AIC). In PCA pre-processing, we used kernel PCA as the principal component analysis. In this method, principal components of 128 features were used as regressor candidates during feature selection, and 42 PCs were used as regressors in multiple regression. Figure 2 suggests a positive linear relationship between HDS-R and SPCIR. The results indicate a moderately significant correlation ($R = 0.781$) between the HDS-R score and the appropriate synthesis of several selected prosodic features. Consequently, the adjusted coefficient of determination ($\bar{R}^2 = 0.54$) suggests that the prosody-based speech sound analysis could potentially be used to detect cognitive impairment in elderly patients.

IV. Classification of NC, MCI, AD Groups using fNIRS signals

The section describes a Bayesian classifier using fNIRS signals of elderlies during cognitive tests, which can discriminate among elderly individuals with three clinical groups: normal cognitive abilities (NC), patients with mild cognitive impairment (MCI), and Alzheimer's disease (AD). To design an algorithm for computer-aided diagnosis of cognitive impairment in the elderly, we consider the screening process by a specialist in geriatrics. We thus propose a two-phase Bayesian classifier shown in Figure 3 on the assumption of screening process, that firstly checks the suspicion of the cognitive impairment (CI) or not (NC) from given fNIRS signals; if any, and then secondly judges the degree of the impairment: MCI or AD.

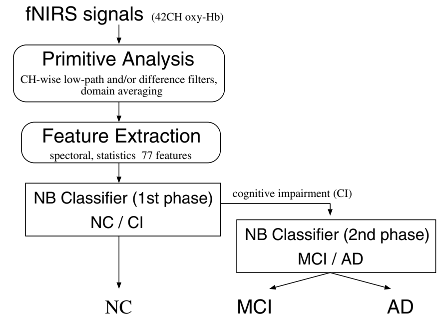

Fig. 3. Classification of NC/MCI/AD by two-phase Bayesian Classifier.

TABLE III

Classification Results (SPCIR Cutoff: 26)

Clinical \ Detection	NC*	MCI	AD	accuracy
NC	19(18)	1	0	95.0%
MCI	6(6)	13	0	68.4%
AD	0(0)	0	9	100%
predictive value	76.0%	92.9%	100%	85.4%

* Value in bracket means the included number of subjects which are classified NC by SPCIR (1st screening).

TABLE IV

Classification Results (SPCIR Cutoff: 28)

Clinical \ Detection	NC*	MCI	AD	accuracy
NC	14(0)	4	2	70.0%
MCI	0(0)	18	1	94.7%
AD	0(0)	1	8	88.9%
predictive value	100%	78.3%	72.7%	83.3%

* Value in bracket means the included number of subjects which are classified NC by SPCIR (1st screening).

A. Primitive Analysis of fNIRS signals

In advance of Bayesian classification, we make a primitive signal processing fNIRS signals. Firstly, we make five fNIRS signals every channels such that noise is reduced by channel-wise smoothing through three low-pass filters and difference filters. Secondly, we segregate 42 CHs into the seven brain areas (see Fr, Fc, Fl, Rf, Rr, Lf, and Lr in Figure 1) and then make signal averaging that integrates fNIRS signals within each of the areas.

B. Extraction of fNIRS Features

We enumerate features that represent fluctuations of regional cerebral blood flow if it is the slightest effective in detection of cognitive impairment, and extract 11 features from fNIRS signals in each of the seven brain areas.

C. Bayesian Classifier

In this paper, we adopted naive Bayes classifier (NB), [14] which is a simple Bayesian classifier with strong independence assumption of attributes. We construct two classifiers: $\text{NB}_{\text{NC/CI}}$, which checks the suspicion of the cognitive impairment (CI) or not (NC) at the first phase, if any suspicion, and $\text{NB}_{\text{MCI/AD}}$, which judges the degree of the impairment (MCI or AD) at the second phase.

V. Classification Assessment

In this paper, we discuss the effectiveness of screening system for dementia and cognitive impairment in elderlies. Our previous work has not discussed a prosody-CBF hybrid screening system, while we have already confirmed the potentiality for either of speech prosody (SPCIR) or CBF (fNIRS). The section describes the prosody-CBF hybrid screening system, which is carried out in two phase: detecting elderly with suspicion of the cognitive impairment (CI) by cut-off with SPCIR score at the first phase, and classifying elderly individuals with three clinical groups NC/MCI/AD using Bayesian classifier at the second phase.

A. Classification Results

We have examined discrimination performance of prosody-CBF hybrid screening system for discriminating among elderly individuals with NC, MCI, and AD, by using simultaneous voice-fNIRS measurement during cognitive tests (1. category recall) collected from 48 participants (see Table I). To evaluate classification performance, we conducted leave-one-out cross-validation.

Table III-IV show the confusion matrices and the statistics of classification results by prosody-CBF hybrid screening, where the cutoff point at the first screening is 26 and 28, respectively. Table III suggests that this cutoff point reduces the cost of fNIRS measurement by half, because the half number (24 participants) of elderlies are classify into NC group at the first screening. At a glance, the system seems to have an acceptable performance, since the accuracy rates of NC and AD were 95% and 100%, respectively, and the total accuracy rate was 85.4%. However, 6 patients with MCI (32% of all participants) were misclassified into NC group. This misclassification should not be negligible for cost-cutting. So, this cutoff point is too sharp, in this particular clinical samples. On the other hand, Table IV suggests that this cutoff point does not reduce the cost of fNIRS measurement at all, because all elderlies are transferred into the second screening. However, the system has an acceptable performance, since the accuracy rates of MCI and AD was 95% and 89%, respectively, and could predict NC perfectly. With cutoff point of 27, system performance were moderate between the above two results. This means a trade-off between cost-cutting and misclassification.

VI. Conclusion

We developed a new technology for early detection of cognitive impairment in the elderly, focusing on the speech prosody and the brain activity during cognitive task. In our previous research, we have studied two data-mining approaches by focusing on speech-prosody and cerebral blood flow (CBF) activation during cognitive tests. On the power of these research results, this paper presented a prosody-CBF hybrid screening system of the elderly with cognitive impairment based on a Bayesian approach. This paper also reported the examination of the detection performance by cross-validation, and the results that both the accuracy rate of AD and the predictive value of NC are equal to or more than 95%. Consequently, the empirical results suggested that proposed approach is adequately practical to screen the elderly with cognitive impairment.

In future work, we will dedicate to the improvement of hybrid methodology of speech-prosody and CBF approaches, and the improvement of SPCIR laying weight on the discrimination of NC/CI. We will also make a follow-up of participants, especially who are classified into MCI, whether he/she will convert to AD or not, and attempt to develop a classification model of convert or not.

Acknowledgments

We are grateful to SHIMADZU Corporation, National Center for Geriatrics and Gerontology, and Ifcom Co. Ltd. for fNIRS measurement system, clinical data collection environment, and data collection, respectively.

References

[1] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ""Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician," *J. Psychiat. Res*, vol. 12, no. 3, pp. 189–198, 1975.

[2] Y. Imai and K. Hasegawa, "The revised Hasegawa's Dementia Scale (HDS-R): evaluation of its usefulness as a screening test for dementia," *J. Hong Kong Coll. Psychiatr.*, vol. 4, no. SP2, pp. 20–24, 1994.

[3] J. C. Morris, "The Clinical Dementia Rating (CDR): Current version and scoring rules," *Neurology*, vol. 43, no. 11, pp. 2412–2414, 1993.

[4] H. Buschke, G. Kuslansky, M. Katz, W. F. Stewart, M. J. Sliwinski, H. M. Eckholdt, and R. B. Lipton, "Screening for dementia with the Memory Impairment Screen," *Neurology*, vol. 52, no. 2, pp. 231–238, 1999.

[5] M. J. de Leon, S. DeSanti, R. Zinkowski, P. D. Mehta, D. Pratico, S. Segal, C. Clark, D. Kerkman, J. DeBernardis, J. Li, L. Lair, B. Reisberg, W. Tsui, and H. Rusinek, "Mri and csf studies in the early diagnosis of alzheimer's disease." *Journal of Internal Medicine*, vol. 256, no. 3, pp. 205–223, 2004.

[6] L. Mosconi, V. Berti, L. Glodzik, A. Pupi, S. De Santi, and M. J. de Leon, "Pre-clinical detection of Alzheimer's disease using FDG-PET, with or without amyloid imaging," *Journal of Alzheimer's Disease*, vol. 20, no. 3, pp. 843–854, 2010.

[7] M. J. de Leon, L. Mosconi, K. B. S. De Santi, R. Zinkowski, P. D. Mehta, D. Pratico, W. Tsui, L. A. Saint Louis, L. Sobanska, M. Brys, Y. Li, K. Rich, J. Rinne, and H. Rusinek, "Imaging and CSF studies in the preclinical diagnosis of Alzheimer's disease," *Annals of New York Academy of Sciences*, vol. 1097, pp. 114–145, 2007.

[8] S. Kato, Y. Suzuki, A. Kobayashi, T. Kojima, H. Itoh, and A. Homma, "Statistical analysis of the signal and prosodic sign of cognitive impairment in elderly-speech: a preliminary study," in *Proc. of 5th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2011)*, 2011, pp. 322–327.

[9] S. Kato, H. Endo, and Y. Suzuki, "Bayesian-Based Early Detection of Cognitive Impairment in Elderly Using fNIRS Signals during Cognitive Tests," in *Proc. of 6th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2012)*, 2012, pp. 118–124.

[10] A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function," *Trends Nurosci.*, vol. 20, pp. 435–442, 1997.

[11] A. Villringer and U. Firnafl, "Coupling of brain activity and cerebral blood flow: basis of functional nuroimaging," *Cerebrovasc. Brain Metab. Rev.*, vol. 7, pp. 240–276, 1995.

[12] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32–80, 2001.

[13] V. Taler, S. R. Baum, H. Chertkow, and D. Saumier, "Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease," *Neuropsychology*, vol. 22, no. 2, pp. 188–195, 2008.

[14] P. Langley, W. Iba, and K. Thompson, "An analysis of Bayesian classifiers," in *Proc. of The Tenth National Conference on Artificial Intelligence (AAAI-92)*, 1992, pp. 223–228.