International Journal of Advanced Research in Science Management and Technology
ISSN(Online)-2454-4159 Volume 3, Issue 6, June 2017

"A Novel on Speech Recognition with Hidden Markov Model"

Patil Priyanka[1], Prof. Mrs. J. H. Patil[2]

PG Scholar, Department of E&TC Engineering, D. N. Patel College of Engineering, Shahada, Maharashtra, India[1]

Professor, Department of E&TC Engineering, D. N. Patel College of Engineering, Shahada, Maharashtra, India[2]

Abstract— The utilization of hidden Markov models (HMMs) has discovered broad use in numerous distinctive zones. This part concentrates on HMMs connected to the performance evaluation of PC frameworks and networks. In the wake of exhibiting a brief audit of foundation material on HMMs, applications, for example, channel postponement and misfortune qualities, movement demonstrating what's more, workload era are overviewed. The force of HMMs as indicators of performance measurements is likewise highlighted. We finish up by introducing a couple components of the module of the Tangram-II performance evaluation device that is focused to HMMs.

Keywords— Hidden Markov models, performance evaluation, network applications.

I. Introduction

Hidden Markov models (HMM) have been utilized as a part of a heap of applications that incorporate discourse acknowledgment, signal preparing, counterfeit consciousness, computational science, money, picture handling, and restorative determination, to give some examples. Reference [2] gives a thorough book reference of more than 300 articles on these applications. The main major fruitful use of HMMs was in discourse acknowledgment [13] and was trailed by application in a far reaching set of use regions. The HMM structure gives a rich hypothetical premise that can be adjusted to a wide range of applications. In spite of the vast assemblage of writing on HMM applications, its utilization in performance evaluation and PC correspondence is moderately new. The presentation of the application of HMMs in performance displaying persuades the present part. In performance displaying, the expert generally has an unmistakable comprehension of how the framework to be investigated works. In the event that a Markovian model is the worldview of decision, the fundamental errand is to choose the fitting framework state variables and the scope of conceivable qualities for each. Case in point, on the off chance that the framework can be demonstrated by a solitary server line and the objective is to anticipate the holding up time for a given landing rate, the state variable of decision is the quantity of clients lined for administration. Other state variables might be presented, for example, to speak to a stage sort administration time appropriation (rather than the exponential dispersion) or to speak to changes in the mean entry rate with time. Another sort of model that falls in this class is accessibility demonstrating.

For this situation, state variables speak to parts that can be operational or in various methods of disappointment. Among the inquiries one may answer is the portion of time the framework is accessible to the client and the likelihood that the framework falls flat in a given time interim.

Typically, models like a straightforward queuing framework or steadfastness model, are parameterized from some earlier information of the framework conduct. A key point to have as a primary concern is that, in these models, one takes the perspective that the framework state is specifically discernible. On the opposite, with HMM the framework condition of the fundamental procedure is accepted to not be specifically discernible. Or maybe one can watch a few values that are a probabilistic capacity of the condition of the basic Markov process. (Hence the starting point of the name Hidden Markov Model).

International Journal of Advanced Research in Science Management and Technology
ISSN(Online)-2454-4159 Volume 3, Issue 6, June 2017

II. LITERATURE SURVEY

In this segment we first compress the documentation to be utilized as a part of whatever remains of this section. In the past segment we depicted a commonplace kind of use of HMMs and the sorts of inquiry that the model might be utilized to reply with regards to that application. Here we compress, autonomous of a specific application, a more formal meaning of a HMM and the most widely recognized classifications of issues crosswise over applications. Later segments will examine specific applications in the setting of performance evaluation. To elucidate these thoughts, consider the accompanying basic illustration. Assume that one is keen on foreseeing the conduct of a client who is skimming through the website pages of an online book store. Expect that we might want to decide the likelihood that a particular client is prepared to arrange a thing taking into account their past conduct inside their session. Assume that, in the wake of considering the conduct of numerous clients, the examiner chooses to manufacture a basic 4-state Markovian model. Expect that the client's states are picked as takes after: simply scanning (JB), looking for things in the store, intrigued by an item (IP) and prepared to arrange a thing (RO). For accommodation we additionally present a leaving (LV) state to the model to demonstrate that a client can leave in the wake of being in any of the past states. (A move to state LV to compares to a session finishing and a new session begins in this particular case with a move to state JB.)

To gauge the move probabilities, we may for instance, take after the direction of a set of clients, where the clients show along the way their present condition of plan. At that point, with this preparation information, one may gauge the state move probabilities from this case information utilizing the relative recurrence of moves from the picked states. Figure 1(a) represents one conceivable Markov chain and its move probabilities.

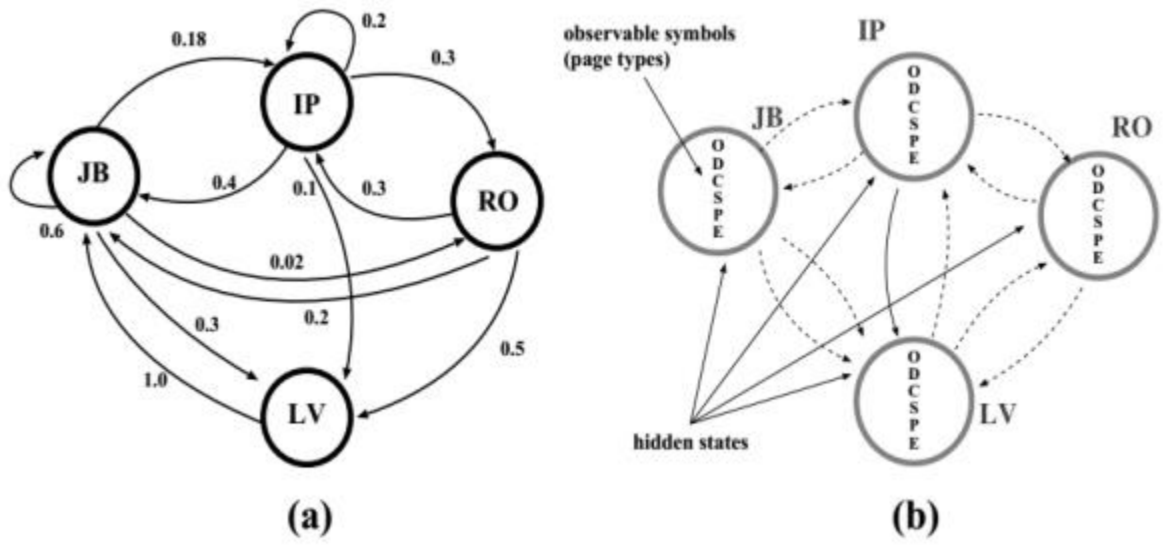


Fig.1 An example of (a) a simple Markovian model of customer behavior; (b) an application example of HMMs.

III. PROBLEM DEFINITION

a client visits and not their condition of goal. For instance, expect that there are 7 distinctive sorts of pages in the site: item diagram (O), item subtle elements (D), set of items inside a classification (C), shopping basket (S), buy (P), exit (E). It is anything but difficult to record the sort of pages the client taps on, however we can't know the client's goal specifically. As an outcome, we can't straightforwardly watch the condition of the client's goal (JB, IP, RO). In any case, it is natural that the states are related with the grouping of perceptions. At the end of the day, theMarkov states JB, IP, RO are hidden in the preparation information yet the grouping of page sorts being gone by is known also, gives some proof of what the client's condition of goal is.

Accept that, some way or another, we can decide the move probabilities of the hidden chain and, for each hidden state, the restrictive likelihood of radiating a discernible image. (In this case, the likelihood of tapping on a specific page sort (O, D, and so on.). Figure 1(b) demonstrates the hidden states and discernible images of the model. A run of the mill inquiry to be asked is what is the most likely hidden state (that is, the client plan) given the watched arrangement of pages went to up to the present time.

Be that as it may, as said above, when all is said in done we don't have the foggiest idea about from the earlier the move probabilities of the hidden chain. Truth be told, in many issues, we don't know what number of states the hidden chain contains. (Envision that you not just can't play out a test, as portrayed prior, where an arrangement of clients intentionally give the data of their purpose yet you may not indeed, even from the earlier know the quantity of or translation of the "conditions of expectation".) For this situation, you know minimal about the hidden MC and can just record the succession of page snaps for various sessions. In any case, to answer the topic of the past section we should first form a totally indicated hidden MC (counting all move probabilities) furthermore the likelihood of clicking a page of every sort molded on the present condition of the hidden MC. The main data accessible might be the grouping of page snaps. For this situation, the issue is to locate the obscure model parameters that boost the likelihood that the watched succession is produced by the model. (As a rule this will rely on upon (an) a from the earlier likelihood dispersion over the space of conceivable models and (b) the contingent likelihood of the watched sequence(s) given a particular model.

IV. RESULT ANALYSIS

Training: The most difficult task is to adjust the model parameter to accurately represent the word under consideration. In training mode large amount of voice data (from different speaker) is given to HMM model. Using this, HMM adjust its probability distribution and transition matrix. There is no global optimal algorithm for learning. Every HMM must be trained to maximize it's (local optimum) recognition power.

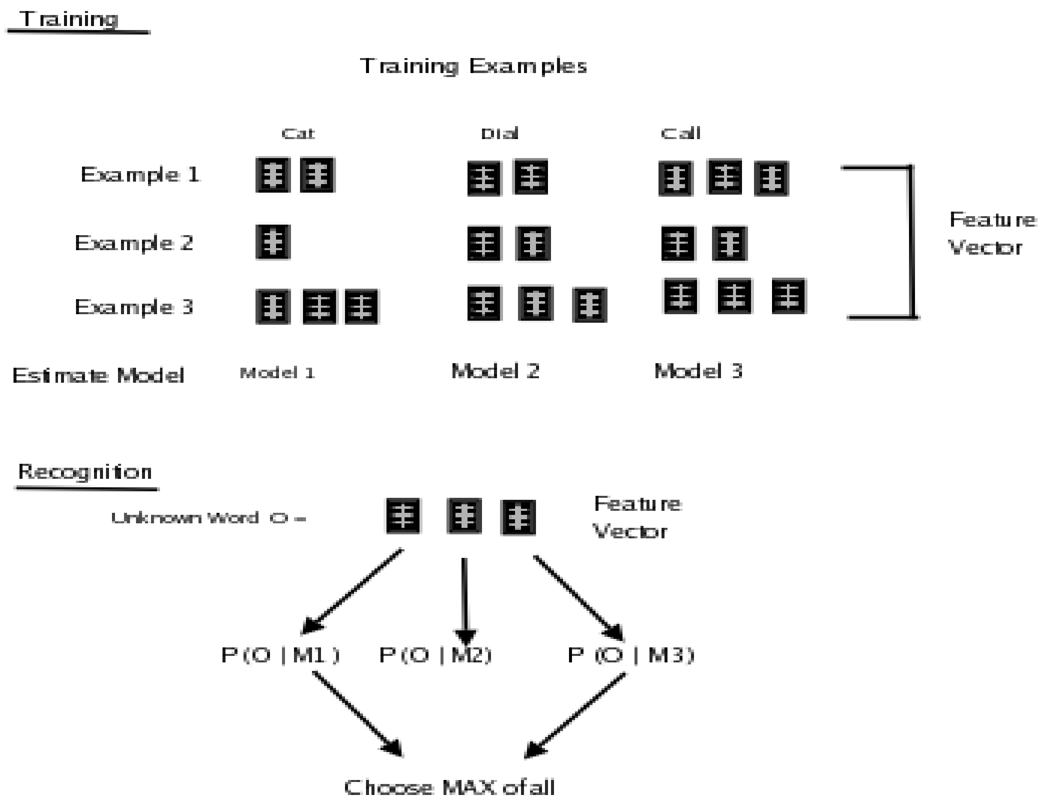


Fig.2 Training Data

I have taken total 9 data set of different voices that is traningdata.then 3 training sets have taken that are compare each every sets that will be get selected.

Fig.3 Create audio data group

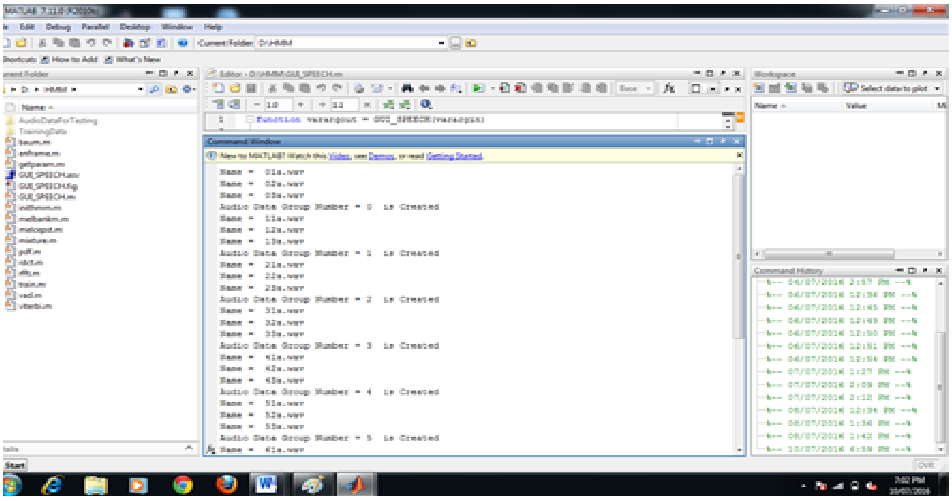


Fig.4 Train data using HMM

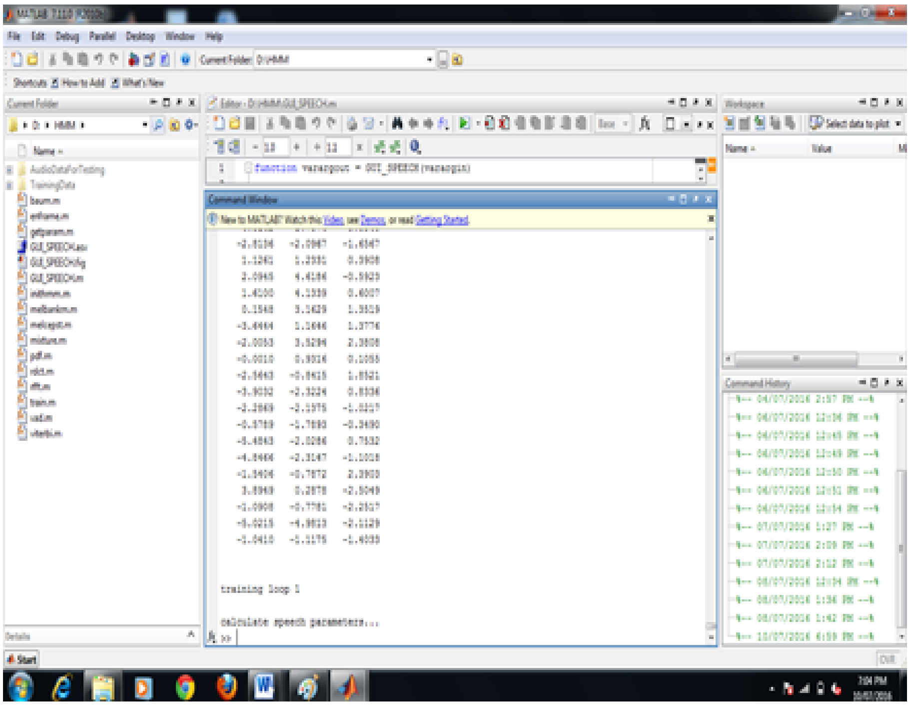
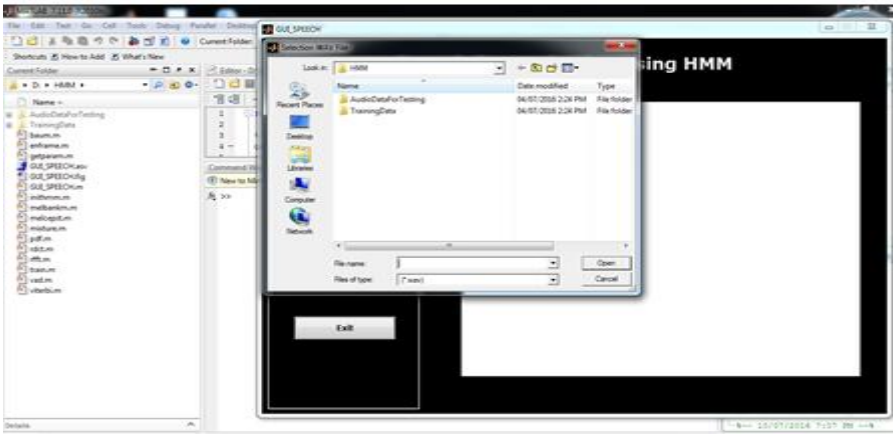


Fig.5 Play audio file for testing

International Journal of Advanced Research in Science Management and Technology
ISSN(Online)-2454-4159 *Volume 3, Issue 6, June 2017*

Initially HMM for phoneme (beforelearning) consists of 3-state and it's output probability distribution areinitialized randomly. It gets automatically updated once the training starts.



That are total 9 data sets of different voice

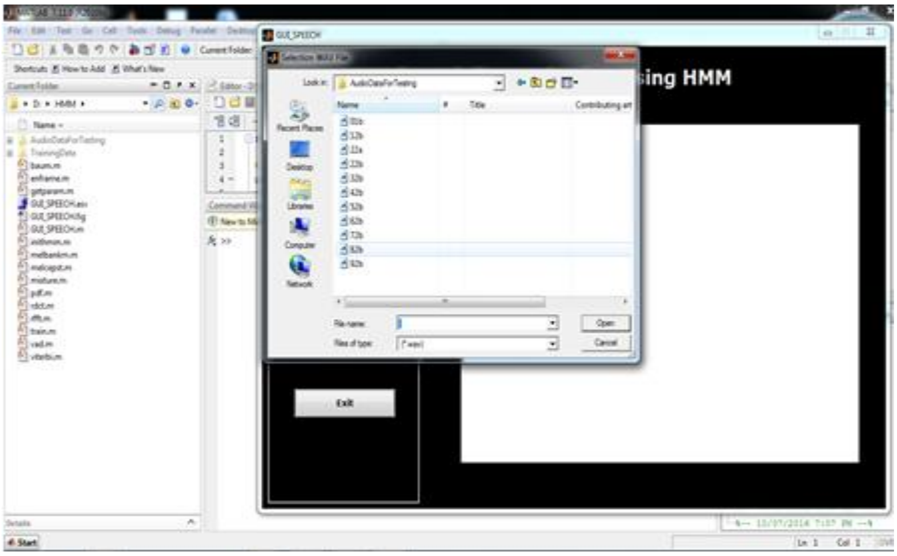


Fig.6 Speech recognition using HMM

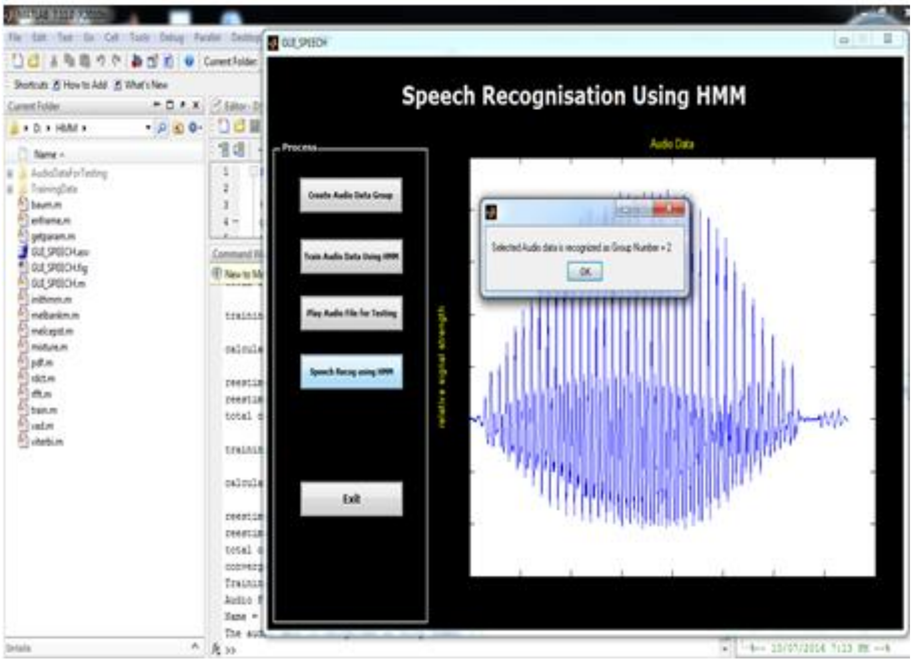


Fig.7 Output of HMM

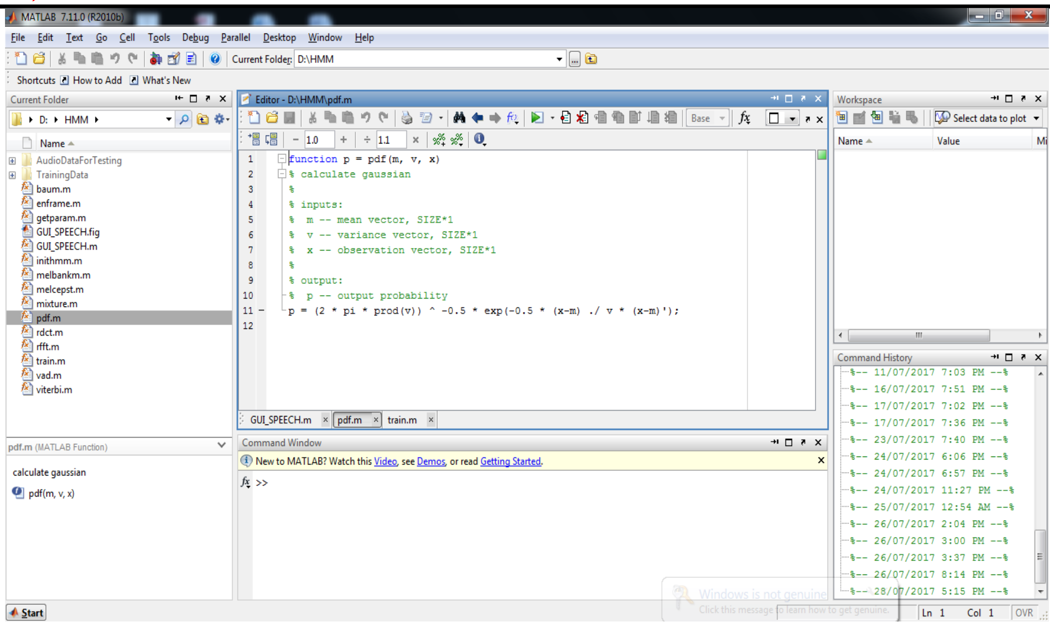


Fig.8 Image of probability

The probability Output

Table no.3.5 The Output probability

Training Loops Steps	Total Output Probability(log)
1.Loops 6	-5.95e+003
2.Loops 13	-4.83e+003
3.Loops 13	-5.38e+003
4.Loops 12	-5.07e+003

V. CONCLUSION

In this work, an EMG-based ASR scheme is proposed and the performance of the proposed scheme was evaluated. Our interests are mainly focused on the representation of the state observation probabilities. The underlying assumption is that inter channel dependencies exist among the features derived from the multichannel EMG data. The dependencies were represented using cross correlational probabilities between individual channel data and global control variables. Several related issues including model parameter estimation were also proposed, based on a maximum likelihood criterion.

The findings here show that HMM derived from the dependent model produced better recognition accuracy than the in-dependent model. This was confirmed by experimental results, yielding up to an 85.07% recognition accuracy where recognition was performed on the isolated words. The resultant recognition accuracy of the proposed model is comparable with that of the early integration model. The proposed approach can be applied to the other multichannel ASR schemes, such as audio visual ASR systems.

To increase the usefulness of the EMG based ASR system, some practical aspects should be considered. For example, the number of words to be recognized should be increased. More-over, the system can recognize more complicated sentences. Another weak point of the proposed scheme is that the locations of electrodes were not determined optimally. Hence, a detailed analysis of the relationship between the function of each facial muscle and the phonemes spoken would be desirable. This will provide clues for determining the optimal locations for collecting the surface EMG signals. Our future studies will focus on these issues.

REFERENCES

[1] S. Kumar, D. K Kumar, M. Alemu, and M. Burry, "EMG based voice recognition," in Proc. IEEE ISSNIP, 2004, pp. 597-596.

[2] H. Manabe and Z. Zhang, "Multi-stream HMM for EMG-based speech recognition," in Proc. 26th Ann. Int. Con. IEEE EMBS, San Francisco, CA, 2004, pp. 4389-4397.

[3] D. Farina, C. Fevotte, C. Doncarli, and R. Merletti, "Blind separation of linear Instantaneous mixtures of nonstationary surface myoelectic signals," IEEE Trans. Biomed. Eng., vol. 51, no. 9, pp. 1555-1567, Sep.2004.

[4] K Ogino and W. M. Kozak, "Spectrum analysis of surface electromygram," in Proc. IEEE Int Conf. Acoustic, Speech, Signal Processing, Boston, MA, 1983, pp. 1114

[5] F. Grandori, P. Pinelli, P. Ravazzani, F. Ceriani, G. Miscio, F. Pisano, R. Colombo, Insalaco, and G. Tognola, "Multiparametric analysis of speech production mechanisms," IEEE Eng. Med. Biol., vol. 4-5, no. 2, pp. 203-209, Apr./May 1994.

[6] H.-J. Park, S.-H.Kwon, H.-C.Kim, and K-S. Park, "Adaptive EMG-driven communication for the disability," in Proc. 1st Joint BMES/ EMBS Con! Serving Humanity, Advancing Technology,Atlanta, GA, 1999, p. 656.

[7] A. D. C. Chan, K B. Englehart, B. Hudgins, and D. F. Lovely, "Hidden Markov model classification of myoelectics signals in speech," IEEE Eng. Med. Biol., vol. 9-10, no.5, pp. 143- 146, Sep./Oct. 2002.

[8] A. D. C. Chan, K B. Englehart, B. Hudgins, and D. F. Lovely, "Myoelectric signals to augment speech recogntion," Med. Bioi. Eng. Comput.,vol. 39, no. 4, pp. 500-504,2001.

[9] A. D. C. Chan, K B. Englehart, B. Hudgins, and D. F. Lovely, "Mul-tiexpert automatic Speech recognition using acoustic and myoelectric signals," IEEE Trans. Biomed. Eng., vol. 53, no. 4, pp.676-685, Apr. 2006.

[10] K-S. Lee, "HMM-based automatic speech recognition using EMG sig-nals," 1.Biomed. Eng. Res. (J-BME), vol. 27, no. 3, pp. 101-109, Jun. 2006.

[11] R. S. Kumaran, K Narayanan, and J. N. Gowdy, "Myoelectric signals for multimodal speech recognition," in Proc. INTERPEECH (Eurospeech), Lisboa, Portugal, 2005, pp. 1189-1192.

[12] A. Q. Summerfield, "Lipreading and audio-visual speech perception,"

[13] L. R. Rabiner and R. W. Schafer, Digital Processing of Speech Signals. Englewood Cliffs, NJ: Prentice-Hall, 1978.

[14] L. R Rabiner and B. H. Juang, Fundamentals of Speech Recognition. Englewood Cliffs, NJ: PrenticeHall, 1993.

[15] G. M. White and R B Neely, "Speech recognition experiments with linear prediction, bandpass filtering, and dynamic programming," IEEE Trans. Acoust. Speech Signal Process.,vol. ASSP-24, no. 2, pp. 183-188, Apr. 1976.

[16] C. Jorgensen, D. D. Lee, and S. Agabon, "Sub auditory speech recognition based on EMG signals," in Proc. Int. Joint Con! Neural Netw.,2003, vol. 4, pp. 3128-3133.

[17] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from in-complete data via the EM algorithm," J. Royal Stat. Soc., vol. 39, pp. 1-38, 1977.

[18] Y. Linde, A. Buzo, and R M. Gray, "An algorithm for vector quantizer design," IEEE Trans .. Communications, vol. 28, no. 1, pp. 84-95, Jan. 1980.

[19] D. A. Reynolds and R C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," IEEE Trans Acoust. Speech Signal Process.,vol. 3, no. 1, pp. 72-83, Jan. 1995.

[20] Blankenstein, B., "Artificial Neural Networks (ANN) ",CS 527A Lecture Notes, http://www.cs.wustl.edu/~sg/CS527_SP01/diagram7_2.gi f, Department of Computer Science & Engineering, Washington University in St. Louis 2001.

[21] Bourlard, H. A. and Morgan, N., ``Connectionist Speech Recognition: A Hybrid Approach'', Kluwer Academic Publishers, Massachusettes, 1994.

[22] Bourlard. H., Konig, Y. And Morgan, N., "REMAP: Recursive Estimation And Maximization Of A Posteriori Propabilities In Connectionist Speech Recognition", Internal report of ICSI, TR-94-064, 1995

[23] Bourlard, H., Dupont, S., Hermansky, H. and Morgan, N., "Towards Subband-based Speech Recognition", Proc. VIII European Signal Processing Conference EUSIPCO'96, Trieste, Italy, 1996.

[24] Chetouani, M., Faúndez-Zanuy, M., Gas, B. and Zarader, J. L., "Non-Linear Speech Feature Extraction for Phoneme Classification And Speaker Recognition", http://www.iiass.it/school2004/School_Materials_1/oral_contributions/Chetouani_short_slides.pdf, Nonlinear Speech Processing School, Vietrisul Mare (Salerno), ITALY, 2004.

[25] Frisch, J., "Modular Neural Networks For Speech Recognition", Master's Thesis, Department Of Computer Science, Carnegie-Mellon University, Pittsburgh Pa, 1996.

[26] Franco, H., Weiniruub, M. and Cohen, M., "Context Modeling in a Hybrid HMM-Neural Net Speech recognition System", Proceedings of the International Conference on Neural Networks, Houston, TX, 1997

[27] Haykin, S., "Neural Networks A Comprehensive Approach", Prentice Hall, 1999

[28] Honkela, T., "Self-Organizing Maps In Natural Langu age Processing", Thesis for a degree of PhD., Neural Networks Research Centre, Helsinki University of Technology, 1997.

[29] Hosom, J. P., Cole, R., Fanty, M., "Speech Recognit ion Using Neural Networks at the Center for SpokenLanguage Understanding", http://speech.bme.ogi.edu/tutordemos/nnet_recog/recog.h tml, OGI School of Science and Engineering, Oregon Graduate Institute of Science and Technology, 1999

[30] Hyvärinen, A.," Survey on Independent Component Analysis", National Computer Symposium 1999, http://www.cis.hut.fi/aapo/papers/NCS99web/, 1999

[31] Jurafsky, D. and Martin, J. H., "Speech AndLanguag e Processing: An Introduction To Natural Language Processing, Computational Linguistics, And Speech Recognition", Prentice Hall, 2000.

[32] Kwon, O-W., Lee, T-W., " ICA-Based Feature Extracti on for Phoneme Recognition", Interspeech 04, 2004

[33] Robinson, A. J., Cook, G. D., Ellis, D. P. W., Fosler-Lussier, E., Renals, S. J., Williams, D. A. G., "Connectionist Speech Recognition Of Broadcast News ", Speech Communication, 37:27-45, 2000

[34] Salmela, P., "Neural Networks in Speech Recognition ", Tampere University of Technology, Publications 295, 2000.

[35] Tebelskis, J., "Speech Recognition Using Neural Networks", PhD Dissertation, Carnegie Mellon University, 1995.

[36] S. Mahd and T. Azizollah, "Voice command recognition system based on MFCC and VQ algorithm, "World Academy of Science,Engineering and Technology Journal, 2009.

[37] W. Stefan and H. Reinhold, "Approaches to iterative speech feature enhancement and recognition, "IEEE Transaction on Audio,Speech and Language Processing, Vol. No.5, July 2009.

[38] F. Sadaoki, "50 years of progress in speech and speaker recognition research," ECTI Transactions on Computer and Information Technology, Vol. 1. No.2, November 2005.

[39] K. H. Davis,R.Biddulp,andS.Balashek, "Automatic recognition of spoken digit," J. Acoust.Soc.Am., 24(6):637 -642, 1952.

[40] H. F. Olson and H. Belar, " Phonetic Typewriter," J. Acoust.Soc.Am., 28(6): 1072-1081, 1956.