ISSN 1828-6003
Vol. 5 N. 6
November 2010

International Review on Computers and Software (IRECOS)

Contents:

(continued on outside back cover)

Praise Worthy Prize

International Review on Computers and Software (IRECOS)

The *International Review on Computers and Software (IRECOS)* is a publication of the **Praise Worthy Prize S.r.l.**.
The Review is published bimonthly, appearing on the last day of January, March, May, July, September, November.

Published and Printed in Italy by **Praise Worthy Prize S.r.l.,** Naples, November 30, 2010.

International Review on Computers and Software (I.RE.CO.S.), Vol. 5, N. 6
November 2010

Eigenspace-Based MLLR Adaptation Using MCE

Reza Sahraeian[1], Behzad Zamani[2], Ahmad Akbari[2], Ahmad Ayatollahi[1]

Abstract – *This paper considers the problem of rapid and robust speaker adaptation in Automatic Speech Recognition (ASR) systems. We propose an approach using combination of eigenspace-based maximum likelihood linear regression (EMLLR) and Minimum Classification Error (MCE). MCE is used to estimate the coefficients of eigenvoices as an alternative to maximum likelihood. Experimental results on TIMIT database illustrate that using MCE leads to up to 1.9% improvement in phoneme error rate. Moreover, it is shown that increasing the amounts of adaptation data has less effect in proposed method than conventional EMLLR. Furthermore, the results achieved using genetic algorithm (GA) to estimate eigenvoices coefficients are compared to proposed method using MCE. The adaptation results using MCE show better performance than GA.*

***Keywords**: Speaker Adaptation, Eigenspace-based Maximum Likelihood Linear Regression, Minimum Classification Error, Genetic Algorithm*

Nomenclature

μ_s	Adapted mean vector
W_s	MLLR transform for speaker *s*
ξ_{si}	Extended mean vector for peakerindependent model
v_d	*d*-th eigenvector
c_d	*d*-th coefficient
$\gamma_r(t)$	Occupation probability of *r* density at time *t*
o	Observation data
Λ	Current model parameter set
α	Tuning parameter
λ	Hidden Markov Model
I	Number of classes
η	Cluster weighting
β	Learning parameter
$\delta(.)$	*Kronecker* delta function
q_t	HMM state at time *t*
$o_{k,t}$	Observation vector for frame *t* of class *k*
$\mu_{jm}^{(i)}(C)$	Mean vector of the *m*-th Gaussianmixture in the state *j* of the *i*-th HMM which is dependent on vector *C*
$\Sigma_{jm}^{(i)}$	Covariance matrix of the *m*-th Gaussian mixture in state *j* of the *i*-th HMM
$a_{jm}^{(i)}$	Weight of the *m*-th Gaussian mixture in the state *j* of the *i*-th HMM
n	Dimensionality of observation vector

I. Introduction

In speaker-independent (SI) speech recognition systems, recognition accuracy varies considerably from speaker to speaker, and performance may be significantly degraded for outlier speakers such as nonnative talkers. Many techniques that have been used to improve the performance and robustness of speech recognition systems include adapting the speech models to the new speaker. Normalizing the input speech of new speakers to some canonical or prototype representation is considered as another approach. In the hidden Markov model (HMM) speech recognition paradigm, adaptation typically means modifying the speaker-independent model parameters based on limited enrolment data from a new speaker.

Speaker adaptation techniques have developed rapidly in the past few years. With a limited amount of speaker specific information, speaker adaptation tries to improve the recognition accuracy of the adapting speaker. A successful speaker adaptation method can benefit many applications including various voice-control appliances, computer aided language learning, dictation software and so on.

Existing speaker adaptation methods can be roughly divided into two types: feature based and model-based methods. Vocal-tract normalization (VTLN)[1] is a feature based adaptation method which eliminates speaker variations caused by different vocal tract lengths. The features are normalized so that the effect of the vocal tract is removed. Some feature-based adaptation methods also consider variations due to other speaker characteristics.

Model-based speaker adaptation techniques can be further divided into three categories: Bayesian-based, transformation-based, and eigenspace-based methods. These methods require the adapting speaker to provide certain amount of speech, and they use the information to improve the recognition performance of the adapting speaker. Due to the fact that clients usually do not want to spend too much time on training the system, it is desirable to have an adaptation method which needs very small amount of adaptation data. This leads to the study of rapid speaker adaptation. Mostly, rapid speaker adaptation means adaptation with less than 10 seconds of adaptation data. It turns out that many conventional adaptation methods break down when there are less than 10 seconds of adaptation data.

Speaker adaptation approaches like the speaker-clustering-based methods [2],[3], the Bayesian-based maximum a posteriori (MAP) adaptation [4], and the transformation-based maximum likelihood linear regression (MLLR) adaptation [5] have been popular for many years. Not only are these methods used in adaptation, but also in speaker recognition [6]. However, when there is small amount of adaptation data for the new speaker-only a few seconds-the eigenvoice-based (or eigenspace-based) adaptation method shows better performance in comparison to other methods [7]. The idea in eigenvoice (EV) speaker adaptation is to derive from a diverse set of speakers a small set of basis vectors called eigenvoices using principal component analysis (PCA) (or other basis-deriving algorithms). These eigenvoices are believed to represent different voice characteristics (e.g., gender, age, accent, etc.). Then a new speaker is represented as a linear combination of a few most important eigenvoices, and the eigenvoice weights are usually estimated by maximizing the likelihood of the adaptation data.

Eigenspace-based adaptation methods including eigenvoice (EV) and eigenspace-based MLLR (EMLLR) which target rapid speaker adaptation, where there is only a very limited amount of adaptation data. The reason for having this limitation is, as mentioned in [8], because clients do not want to spend time training the system.

In [9], the authors propose three approaches, known as EMLLR adaptation, which combine MLLR and eigenvoice adaptation. Kernel methods have been investigated as another way to improve the conventional eigenspace-based adaptation methods by exploiting possible nonlinearity in their working space [10]-[11]. Many techniques have been investigated to improve the performance of recognition[12].

Minimum classification error (MCE) is a well-known discriminative method used for both feature transformations and classifier training [13]. Using MCE in adaptation of parameters of Gaussian mixture continuous density HMM was first reported in [14]. In discriminative linear regression adaptation techniques, MCE is used as an alternative to maximum likelihood (ML) to estimate transformation matrix [15].

Inspired by the success of minimum classification error as an alternative to Expectation Maximization (EM) in maximum likelihood estimation, we propose to use MCE in the framework of eigenspace-based speaker adaptation (MCE-EMLLR). Moreover, MCE is an iterative algorithm provides us with better solution in each iteration which can overcome the drawback of expectation maximization algorithms in taking the global optimum as a final solution. In our recent work, furthermore, using genetic algorithm is taken into the consideration as another approach to estimate eigenvoices coefficients (GA-EMLLR)[16]. Since, genetic algorithm is not sensitive to the amount of adaptation data, GA-EMLLR method is expected to outperforms conventional EMLLR approach using maximum likelihood. Experimental results show that using MCE yields better performance than GA. However, we should contemplate the point that both genetic algorithm and minimum classification error are iterative and slow. Thus, our proposed approaches is mostly applicable in offline recognition.

The use of genetic algorithm [17] in speaker adaptation was first proposed to estimate new speaker model parameters [18]. In [18], Genetic Algorithms have been used to enrich the set of SD systems generated by the eigen-decomposition. Besides, evolutionary based linear transform in speaker adaptation is proposed to estimate the parameters of MLLR transforms for new speaker [19].

The remainder of this paper is organized as follows. Section II presents the principles of eigenspace-based MLLR adaptation. Section III describes the MCE formulation used in our eigenspace-based MLLR adaptation paradigm. Section IV explains the characteristics of evolutionary EMLLR method briefly. Section V is devoted to experimental results of our proposed approach on TIMIT database. Finally, this paper is concluded in section VI.

II. Eigenspace-Based Maximum Likelihood Linear Regression

In standard eigenvoice speaker adaptation [6], a set of speaker-dependent acoustic models are estimated from speech data collected from many training speakers with diverse speaking or voicing characteristics. However, preparing enough data to train speaker dependent system is not easy. Using MLLR transformation matrix instead of a model for each speaker is known as eigenspace-based MLLR adaptation.

Suppose there are some speech data from *N* speakers and an *SI* model that is a hidden Markov model (HMM). For the sake of simplicity, one transformation matrix is considered for each speaker. Consider μ_s as an adapted mean vector:

$$\mu_s = W_s \xi_{si} \tag{1}$$

$\xi_{si} = \left[\mu^{(si)'}, 1\right]'$ is the extended mean vector of corresponding Gaussians in SI models.

In EMLLR adaptation, a speaker is indirectly represented by a speaker transformation supervector (STSV) which is obtained by stacking up the vectorized MLLR transformation matrices:

$$STSV = Y = \{y_1, ..., y_i, ..., y_s\} \tag{2}$$

where $y_i = Vec(W_i)$.

To obtain eigenvectors, principle component analysis (PCA) is preformed using correlation matrix. STSV is mean-zeroed and then normalized by its variance. For a new speaker, the centered and normalized speaker transformation supervector $\hat{y}$ is approximated as a linear combination of the leading vectorized eigenvoices:

$$\hat{y} = \sum_{d=1}^{D} c_d v_d \tag{3}$$

If these coefficients are estimated using maximum likelihood estimation, the following set of equations should be solved:

$$\begin{aligned} &\sum_t \sum_r \gamma_r(t)\left(o_t - \bar{W}\xi_r\right)^T \Sigma_r^{-1} W^{(i)} \xi_r = \\ &= \sum_t \sum_r \gamma_r(t) \times \left\{ \sum_{d=1}^{D} c_d W^{(d)} \xi_r \right\} \Sigma_r^{-1} W^{(i)} \xi_r \end{aligned} \tag{4}$$

$W^{(i)}$ is the i-th eigenmatrix and $\bar{W}$ is the speaker mean MLLR matrix over all speakers.[9].

III. MCE Eigenspace-Based MLLR Formulation

In the MCE framework formulation a sigmoid function is adopted to approximate the empirical error of observation of class k, O_k, under the current model parameter set, Λ, as follows:

$$l_k(O_k, \Lambda) = \frac{1}{1 + exp\left(-\alpha d_k(O_k, \Lambda)\right)} \tag{5}$$

in the above equation, $d_k(O_k, F)$ is a cost function (or misclassification measure) defined as follows:

$$\begin{aligned} d_k(O_k, \Lambda) &= -g_k(O_k, \Lambda) + \\ &+ \frac{1}{\eta} log\left[\frac{1}{I-1} \sum_{i=1, i\neq k}^{I} exp\left(g_i(O_k, \Lambda)\eta\right)\right] \end{aligned} \tag{6}$$

η is tuned empirically and is a positive number represented in [20]. $g_i(O_k, \Lambda)$ is a discriminant function which is logarithm likelihood of observation, O_k, defined as:

$$g_i(O_k, \Lambda) = log\ p(O_k, Q; \lambda_i) \tag{7}$$

$p(O_k, Q; \lambda_i)$ is likelihood for O_k by Hidden Markov Model λ_i with optimal state sequence Q. Total classification error is defined as follows:

$$L = \sum_{k=1}^{I} l_k(O_k, \Lambda) \tag{8}$$

We want to find the eigenvioces coefficients which minimize L. Considering $C = (c_1, c_2, ..., c_K)^T$ as a vector of eigenvoices coefficients; we use gradient decent method to find the optimum solution for these coefficients as represented in [21]:

$$C_{iter} = C_{iter-1} - \beta \frac{\partial L}{\partial C} \tag{9}$$

iter denotes iteration number in gradient descent algorithm. To calculate derivative of function L with respect to C, we present the following formulation:

$$\frac{\partial l_k(O_k, \Lambda)}{\partial C} = \frac{\partial l_k(O_k, \Lambda)}{\partial d_k(O_k, \Lambda)} \frac{\partial d_k(O_k, \Lambda)}{\partial C} \tag{10}$$

$$\frac{\partial l_k(O_k, \Lambda)}{\partial d_k(O_k, \Lambda)} = \alpha l_k(O_k, \Lambda)\left(1 - l_k(O_k, \Lambda)\right) \tag{11}$$

To compute the term $\frac{\partial d_k(O_{n_k}, \Lambda)}{\partial C}$, we use (6):

$$\begin{aligned} \frac{\partial d_k(O_k, \Lambda)}{\partial C} &= -\frac{\partial g_k(O_k, \Lambda)}{\partial C} + \\ &+ \sum_{i=1, i\neq k}^{I} \left\{ \frac{exp\left(g_i(O_k, \Lambda)\eta\right)}{\sum_{j=1, j\neq k}^{I} exp\left(g_j(O_k, \Lambda)\eta\right)} \times \frac{\partial g_i(O_k, \Lambda)}{\partial C} \right\} \end{aligned} \tag{12}$$

Now, we need to compute $\frac{\partial g_i(O_k, \Lambda)}{\partial W}$. Considering relation (7), we have:

$$\frac{\partial g_i(O_k, \Lambda)}{\partial C} = \sum_{t=1}^{T} \delta(q_t - j) \frac{1}{b_j^{(i)}(o_{k,t})} \frac{\partial b_j^{(i)}(o_{k,t})}{\partial C} \tag{13}$$

T is the number of observations and $b_j^{(i)}(o_{k,t})$ is the generated probability for observation vector $o_{k,t}$ in the j-th state of the i-th HMM, defined as:

$$b_j^{(i)}(o_{k,t}) = \sum_{m=1}^{M} a_{jm}^{(i)} b_{jm}^{(i)}(o_{k,t}) = \frac{1}{(2\pi)^{n/2}} \times \sum_{m=1}^{M} \frac{c_{jm}^{(i)}}{\left|C_{jm}^{(i)}\right|^{1/2}} exp\left(\begin{array}{l} -\frac{1}{2}\left(o_{k,t} - \mu_{jm}^{(i)}(C)\right)^T \left(\Sigma_{jm}^{(i)}\right)^{-1} \times \\ \times \left(o_{k,t} - \mu_{jm}^{(i)}(C)\right) \end{array} \right) \tag{14}$$

Thus, the relation in (13) can be written as follows:

$$\frac{\partial g_i(O_k, \Lambda)}{\partial C} = \sum_{t=1}^{T} \delta(q_t - j) \sum_{m=1}^{M} \gamma_{jm}^{(i)}(o_{k,t}) \left[\begin{array}{c} \left(o_{k,t} - \mu_{jm}^{(i)}(C)\right)^T \cdot \\ \cdot \left(\Sigma_{jm}^{(i)}\right)^{-1} \frac{\partial \mu_{jm}^{(i)}(C)}{\partial C} \end{array} \right] \tag{15}$$

We introduce the following notation to compute $\frac{\partial \mu_{jm}^{(i)}(C)}{\partial C}$:

$$\mu_{jm}^{(i)}(C) = \left(\sum_{d=1}^{D} c_d v_d \right) \xi_{jm}^{(i)} = \sum_{d=1}^{D} c_d v_d \xi_{jm}^{(i)} \tag{16}$$

where $\xi_{jm}^{(i)}$ is the extended mean vector of the m-th Gaussian mixture in the state j of the i-th HMM corresponding to speaker independent system.

Implementation of formulation explained above on TIMIT dataset shows the better performance in EMLLR adaptation. Evaluation of proposed approach on continuous speech recognition is described in section V.

IV. Evolutionary Eigenspace-Based MLLR Adaptation Paradigm

The Genetic algorithm (GA) is a randomized search method based on natural evolution. GA copes with search in complex and large spaces, and usually provides near-optimal solutions for a defined fitness function of an optimization problem.

In GA, each instance of search space is encoded in the form of string called chromosome (genotype or individual). A collection of these chromosomes is called population. The initial population is usually created at random and it contains some points in the search space. Defined fitness function presents the goodness of each chromosome, and based on that, the fitter chromosomes are selected for next population. The variation operators such as mutation and crossover are applied on the chromosomes to yield a new generation of chromosomes. The process of selection, crossover and mutation continue for a fixed number of generations or till the termination criterion is satisfied.

In this paper, we use the following fitness function:

$$f(s) = \frac{1}{-log\left(p(o \mid \lambda_s)\right)} \tag{17}$$

where o and λ_s respectively represent the adaptation data and acoustic models of the solution s. In other words, in each of iteration a new adapted model is achieved for each chromosome. The likelihood of adaptation data is calculated for each new model and a score or fitness value is dedicated for that chromosome. The best chromosome is kept for the new population in the next iteration.

Others parameters in our genetic algorithm are defined as follows.

IV.1. Definition of Chromosomes

The first step in defining a GA is to link the “real world” to the “GA world”, that is to setup a bridge between the original problems context and the problem solving space. The chromosomes in this Genetic algorithm are 1×n vectors, where n is the number of coefficients which should be estimated.

IV.2. Population Initialization

The first population is usually selected from randomly generated chromosomes. However, we also examine the case that one proper solution exists in the first population. The existence of such a solution in first population causes the genetic algorithm to reach better result in a short time. The value of each position in each chromosome is a real number between [-MaxValue, MaxValue]. MaxValue is a parameter of the problem and we initialize it with 1.

IV.3. Crossover Operation

Crossover is a probabilistic process that exchanges information between parent chromosomes for generating child chromosomes. We use uniform crossover as the crossover operation with a fixed crossover probability of μ_C =0.9 [17].

IV.4. Mutation Operation

Four Mutation methods are implemented to increase the performance of the genetic algorithm. Mutation operation is applied on the chromosomes with μ_m probability (μ_m=0.1):

- *Random mutation*: In this mutation, a chromosome is selected at random, and then one bit of the chromosome is revalued randomly. The selection rate of the mutation is 25% of μ_m [17].
- *Swap mutation*: In this mutation, a chromosome is selected at random then two bits of this chromosome are selected randomly and then the values of these bits are swapped. The selection rate of this mutation is 25% of μ_m [17].
- *Creep mutation*: This mutation works by adding a small random value to a bit of the selected chromosome and changing the value of this bit. The small random value is a real number between [-CreepValue, CreepValue]. CreepValue is a parameter of the problem and we initialize it with 10. The selection rate of this mutation is 30% of μ_m [17].
- *Scramble mutation*: This mutation is the most disruptive mutation. A chromosome is selected at random and the values of this chromosome are reconfigured. The selection rate of this mutation is 20% of μ_m [17].

V. Experimental Results

To examine the viability and the efficiency of proposed MCE-EMLLR method, a series of experiments for continuous speech recognition of TIMIT are performed. In these experiments, rapid speaker adaptation using maximum likelihood and minimum classification error are taken into consideration.

In all experiments, we use 39-dimension feature vectors consisting of energy, 12 MFCCs and their first and second order derivatives. The features were normalized to have zero mean and standard deviation one over TIMIT training set. We don't reduce features vector dimension in our methods. In addition, we use HMMs with 3 states and 16 Gaussian mixtures per state.

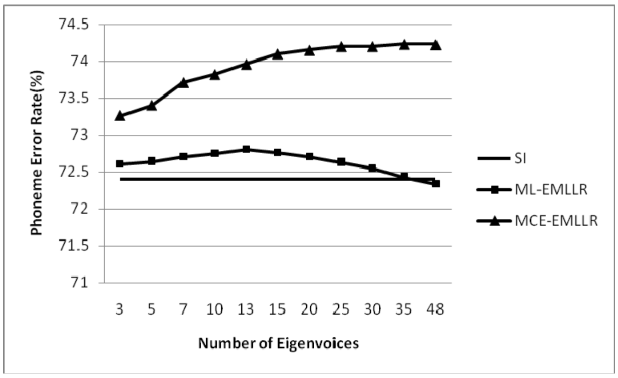

Fig. 1. Phoneme error rate achieved using MCE (MCE-EMLLR) and maximum likelihood (ML-EMLLR) and baseline (SI)

To provide proper set of reference speaker needed for eigenspace-based adaptation, six speakers, three men and three women, from each dialect of TIMIT test set are selected to have total number of 48 speakers; the one over TIMIT training set. We don't reduce features vector dimension in our methods. In addition, we use HMMs with 3 states and 16 Gaussian mixtures per state.

To provide proper set of reference speaker needed for eigenspace-based adaptation, six speakers, three men and three women, from each dialect of TIMIT test set are selected to have total number of 48 speakers; theremainder speakers of test set are used for test. One MLLR transformation matrix is estimated for each of the speakers of reference set. To estimate the coefficients of eigenvoices we conduct our experiments on both maximum likelihood criteria and MCE algorithm as explained in section III. Moreover, we examine the result of EMLLR adaptation using two different amounts of adaptation data. Since eigenvoice adaptation shows good performance in sparse adaptation data, the experiments are done with small speech data.

Fig. 1 represents the results of phoneme error rate in eigenspace based MLLR supervised speaker adaptation. The results obtained for two utterances (about 6 seconds) as an adaptation data. Fig. 1 shows the improvements achieved using minimum classification error algorithm instead of maximum likelihood in eigenspace-based MLLR adaptation.

The experiments are also conducted on two different amounts of adaptation data. Table I represents the achieved results. It is shown that increasing the adaptation speech data leaves more effect on recognition rate than MCE-EMLLR does. In Table I the results of experiments by using 13 eigenvoices are reported.

However, minimum classification error is an iterative algorithm containing parameters which should be set empirically and maybe time consuming. MCE outperforms maximum likelihood because it searches to find better solution in different iterations and amounts of adaptation data has less effect on the result of adaptation.

Furthermore, experimental result achieved using genetic algorithm is given in Table II. In this table the result of phoneme error rate using genetic algorithm is compared to one using MCE.

TABLE I
PHONEME ERROR RATE (%) FOR DIFFERENT AMOUNTS OF ADAPTATION DATA

Different Methods	2 Utterances	3 Utterances
SI	72.4	72.4
ML-EMLLR	72.8	73.12
MCE-EMLLR	73.96	73.97

TABLE II
PHONEME ERROR RATE FOR DIFFERENT METHODS

Different Methods	Phoneme Error Rate (%)
SI	72.4
ML-EMLLR	72.8
MCE-EMLLR	73.96
GA-EMLLR	73.66

Table II reports the results for two utterances (about 6 seconds) as a speech adaptation data when 13 eigenvoices are used. Using genetic algorithm in EMLLR adaptation leads to 1.09% improvements in phoneme error rate. Individuals in first population in GA algorithm used are made randomly. Each population contains 50 chromosomes and GA algorithm is iterated for 80 times.

VI. Conclusion

In this paper the efficiency of minimum classification error to estimate the eigenvoices coefficients in eigenspace-based MLLR adaptation is illustrated. Experiments show the effectiveness of using MCE compared to maximum likelihood in conventional EMLLR adaptation. The proposed adaptation method was tested using TIMIT dataset. Experimental results show that up to 1.9% improvement in phoneme error rate is obtained in proposed method. Experimental results also concerns about the effect of adaptation data available for each speaker. Increasing adaptation data improves the recognition rate; the improvement when maximum likelihood is used is more considerable. Moreover, genetic algorithm as another alternative to maximum likelihood is compared to minimum classification error. MCE outperforms GA for eigenvoices coefficients estimation.

References

[1] E. Eide, H. Gish. A parametric approach to vocal tract length normalization, *In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1*, 1996 , pp. 346–348.

[2] T. Kosaka, S. Matsunaga, and S. Sagayama, Speaker-independent speech recognition based on tree-structured speaker clustering, *Computer Speech and Language, vol. 10*, 1996, pp. 55–74.

[3] T. J. Hazen, A comparison of novel techniques for rapid speaker adaptation, *Speech Communication., vol. 31*, May 2000, pp. 15–33.

[4] J. L. Gauvain, C. H. Lee, Maximum a posteriori estimation for multivariate Gaussian mixture observations of Markov chains, *IEEE Transaction Speech Audio Processing, vol. 2*, April 1994, pp. 291–298.

[5] C. J. Leggetter, P. C. Woodland, Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models, *Computer Speech and Language, vol. 9,* 1995, pp. 171–185.

[6] N. Seman, J. Kamaruzaman, Performance of Adapting Non-Native Speech in Isolated Speech Recognizer, *International Review on Computers and Software, Vol. 3, n. 3*,2008, pp.324-328.

[7] R. Kuhn, J. C. Junqua, P. Nguyen, N. Niedzielski, Rapid Speaker Adaptation in Eigenvoice Space, *IEEE Transaction Speech and Audio Processing, vol. 8,* 2000, pp. 695-707.

[8] R. Kuhn, F. Perronnin, J-Claude Junqua. Time is money: Why very rapid adaptation matters. *In The International Speech Communication Association (ISCA) workshop*, 2001.

[9] N. Wang, S. Lee, F. Seide, L. Lee, Rapid speaker adaptation using a priori knowledge by eigenspace analysis of MLLR parameters, *in Proc. of ICASSP*, 2001, pp. 317–320.

[10] B. Mak, J. T. Kwok, S. Ho, Kernel eigenvoice speaker adaptation, *IEEE Transaction Speech Audio Processing, vol. 13,* September 2005, pp. 984–992.

[11] B. Mak, R. Hsiao, Kernel eigenspace-based MLLR adaptation, IEEE *Transactions Speech and Audio Processing, vol. 15*, March 2007, pp. 784-795.

[12] Z. Sakka, A. Kachouri, M. Samet, Speech Denoinsing and Arabic Speaker Recognition System Using Subband Approach, *International Review on Computers and Software, Vol. 2, n. 3*,2007, pp. 264-271.

[13] B. Zhang, S. Matsoukas, Minimum phoneme error based heteroscedastic linear discriminant analysis for speech recognition, *In Proceedings of ICASSP, vol. 1*, 2005, pp. 925-928.

[14] C.-H. Lin, C.-H. Wu, P.-C. Chang, A study on speaker adaptation for Mandarin syllable recognition with minimum error discriminative training, *IEICE Trans. Inf. & Syst., Vol.E78-D*, No.6, 1995, pp.712-718.

[15] J. Wu, Q. Huo, A study of Minimum Classification error (MCE) Linear regression for supervised adaptation of MCE-trained Continues-Density Hidden Markov Model , *IEEE Trans. On Audio, Speech and Language Processing, vol. 15*, no. 2, Feb. 2007.

[16] R. sahraeian, M. Mohammadi, A. Akbari, A. Ayatollahi, Evolutionary Eigenvoice MLLR speaker adaptation, *Procedia-Computer sience, Elsevier*, 2011.

[17] A. Eiben, J.E. Smith, *Introduction to Evolutionary Computing*. (2nd ed. Springer, 2007).

[18] F. Lauri, I. Illina, D. Fohr, F. Korkmazsky, Using genetic algorithms for rapid speaker adaptation, *in Proc. Eurospeech*, 2003, pp. 1497-1500.

[19] S. A. Selouani, D. O'Shaughnessy, Speaker Adaptation using Evolutionary-Based Linear Transfornation, *in Proc. Interspeech*, Sep. 2006, pp. 1109–1112.

[20] B. H. Juang, W. Chou, C. H. Lee, Minimum classification error rate methods for speech recognition, *IEEE Transaction on Speech and Audio Processing, Vol. 5*, No. 3, 1997, pp. 257-265.

[21] A, De La Torre, A. M. Peinado, A. J. Rubio, V. E. Sanchez, J. E. Diaz, An Application of Minimum Classification Error to Feature Space Transformation for Speech Recognition, *Speech Communication, No. 20*, 1996, pp. 273-290.

Author's information

[1]Electrical Engineering Department, Iran University of Science and Technology, Tehran, Iran.

[2] Computer Engineering Department, Iran University of Science and Technology, Tehran, Iran..

Reza Sahraeian has received his B. Sc. and M.Sc. from Electrical Engineering Department, Iran University of Science & Technology (IUST) in 2006 and 2009 respectively. He is a student member in IEEE since 2008. He is currently a member of Audio and Speech Processing Lab. (ASLP) in IUST. His research interest includes Speech processing and recognition and also discriminative algorithms.

Behzad Zamani has received his M.Sc. from Computer Engineering Department, Iran University of Science & Technology (IUST) under supervision of Dr. Ahmad Akbari. He also has got his B.Sc. from Electronic & Computer Engineering Department of Isfahan University of Technology (IUT).Currently He is a PH.D Condidate in Artificial Intelligence and working as a Research Assistant in the ASPL and RCIT Labs of Computer Engineering Department, Iran University of Science & Technology (IUST).

Ahmad Akbari is an associated professor in Computer Engineering department at IUST. He has received his M.Sc. in communication in Isfahan University of technology in 1989. He got his Ph.D. in Rennes University in signal processing in 1995. He is the head of Audio and Speech processing Lab., and his current research contains speech processing and network security.

Ahmad Ayatollahi is an associated professor in Electrical Engineering department of Iran University of Science and Technology (IUST). He has received his M.Sc. and Ph.D. in Electrical Engineering at UMIST. His current research interest is biometric and image processing.

International Review on Computers and Software (IRECOS)

Aims and scope

The ***International Review on Computers and Software (IRECOS)*** (**www.praiseworthyprize.com/IRECOS.htm**) is a peer-reviewed journal that publishes original theoretical and applied papers on all aspects of Computers and Software. The topics to be covered include, but are not limited to:

technology, hardware systems, parallel and distributed computing, computer graphics (modelling, animations, rendering, visualization, imaging and image processing, human-computer interaction, computer graphics and visualization for environmental decision support; and so on), Internet and multimedia systems (web and internet system and tools, multimedia information systems, data management, collaborative systems and distance learning and so on) computer organizations and architectures; wireless and mobile computing, networking and communication; operating systems, real-time systems and embedded systems; pattern recognition; digital devices, computer components and interconnection networks; specification, design, prototyping and testing methods and tools; performance, fault tolerance, reliability, dependability and secure computing, testability; development and maintenance methods and models, e.g., techniques and principles for the specification, design and implementation of software systems, including notations and process models; assessment methods, e.g., software tests and validation, reliability models, test and diagnosis procedures, software redundancy and design for error control and the measurements and evaluation of various aspects of the process and product; geoinformatics (Web-based geospatial data and GIS services, Location based services (LBS), Automated object extraction and database updates from imagery, Spatial data fusion and integration algorithms, GNSS interoperability, software and algorithms, Integration of environmental modelling and GIS, Wireless GIS solutions for disaster management, Data and systems interoperability for handling emergencies, Visualization of emergency scenes and situations, and so on);software tools and techniques; software project management, e.g., productivity factors, cost models, schedule and organizational issues, standards; tools and environments, e.g., specific tools, integrated tool environments including the associated architectures, databases, and parallel and distributed processing issues; system issues, e.g., hardware-software trade-off; new and important applications and trends.

Instructions for submitting a paper

Contributions may consist of invited tutorials or critical reviews; original scientific research papers (regular paper); letters to the Editor and research notes which should also be original presenting proposals for a new research, reporting on research in progress or discussing the latest scientific results in advanced fields.
All papers will be subjected to a fast editorial process.
Any paper will be published within two months from the submitted date, if it has been accepted.

Papers must be correctly formatted, in order to be published.
Formatting instructions can be found in the last pages of the Review.
An ***Author guidelines*** template file can be found at the following web address:
http://www.praiseworthyprize.com/Template_of_IRECOS.doc

Manuscripts should be sent via e-mail as attachment in .doc and .pdf formats to:
irecos@praiseworthyprize.com or **santolo@unina.it** (Managing Editor)

The regular paper page length limit is defined at **15** formatted Review pages, including illustrations, references and author(s) biographies.
Pages 16 and above are charged 10 euros per page and payment is a prerequisite for publication.

After a paper has been accepted, the author or the author's organization will be requested to buy 4 copies of the Review issue to cover part of the publication cost.

Subscription rates:

	on Cd-Rom, per year:	*Print copy, per year:*
Institutional:	150 (euros)	240 (euros)
Individual:	120 (euros)	210 (euros)
Individual Article:	15 (euros)	25 (euros)

Abstracting and Indexing Information:

Cambridge Scientific Abstracts (CSA/CIG)
Academic Search Complete (EBSCO Information Services)
COMPENDEX - Elsevier Bibliographic Database
Copernicus

Autorizzazione del Tribunale di Napoli n. 59 del 30/06/2006

(continued from outside front cover)

1828-6003(201011)5:6;1-P