IMPROVING EIGENSPACE-BASED FUZZY LOGIC SYSTEM USING A LINEAR INTERPOLATION SCHEME FOR SPEECH PATTERN RECOGNITION

Ing-Jr Ding and Chih-Ta Yen
Department of Electrical Engineering, National Formosa University, Taiwan
E-mail: eugen.ding@gmail.com; chihtayen@gmail.com

ICETI 2012-E2006_SCI
No. 13-CSME-33, E.I.C. Accession 3491

ABSTRACT

The Eigen-FLS approach using an eigenspace-based scheme for fast fuzzy logic system (FLS) establishments has been attempted successfully in speech pattern recognition. However, speech pattern recognition by Eigen-FLS will still encounter a dissatisfactory recognition performance when the collected data for eigen value calculations of the FLS eigenspace is scarce. To tackle this issue, this paper proposes two improved-versioned Eigen-FLS methods, incremental MLED Eigen-FLS and EigenMLLR-like Eigen-FLS, both of which use a linear interpolation scheme for properly adjusting the target speaker's Eigen-FLS model derived from an FLS eigenspace. Developed incremental MLED Eigen-FLS and EigenMLLR-like Eigen-FLS are superior to conventional Eigen-FLS especially in the situation of insufficient data from the target speaker.

Keywords: speech pattern recognition; Eigen-FLS; EigenMLLR-like Eigen-FLS; incremental MLED Eigen-FLS.

AMÉLIORATION DE LA BASE DE L'ESPACE PROPRE UTILSANT LE SYSTÉME DE LA LOGIQUE FLOUE (FLS) DANS UN SCHÉMA D'INTERPOLATION LINÉAIRE POUR LA RECONNAISSANCE DE LA PAROLE

RÉSUMÉ

L'approche Eigen-FLS dans un schéma de base de l'espace propre à systéme de logique floue rapide a été appliquée avec succès pour la reconnaissance de la parole. Toutefois, le modèle de reconnaissance de la parole par Eigen-FLS aura encore une reconnaissance de la parole insatisfaisante quand la collection de données pour la valeur propre des calculs de l'espace propre FLS est aussi peu abondante. Pour s'adresser à cette question, cet article propose deux versions améliorées des méthodes Eigen-FLS, MLED EIGEN-FLS et EigenMLLR comme EIGEN-FLS, les deux utilisant un schéma d'interpolation linéaire pour un ajustement adéquat du modèle EIGEN-FLS du locuteur cible, dérivé de l'espace propre FLS. L'incrémentiel développé MLED Eigen-FLS et EigenMLLR comme espace propre Eigen-FLS sont supérieurs au Eigen-FLS conventionnel, spécialement en situation de données insuffisantes du locuteur cible.

Mots-clés : modèle de reconnaissance de la parole ; Eigen-FLS ; Eigen-MLLR-like EigeN-FLS ; incrémentiel MLED Eigen-FLS.

1. INTRODUCTION

Pattern recognition techniques have been very popular in the recent years and frequently seen in surveillance or security applications, such as community surveillance and access control systems. As is commonly known, video-based (or image-based) pattern recognition, such as hand-written recognition [1], has been an extremely matured technique and widely used in daily life. It has been so for over decades despite the fact that for thousands of years in human history instant and precise communications among individuals were mostly realized via the auditory channels: speak and listen (imagine the age before the creation of characters in ancient civilization). Following such the line of thoughts, audio-based (or speech-based) pattern recognition techniques are rising [2]. Automatic speech recognition (ASR), speaker recognition and acoustic event detection/classification have gradually become more matured and acceptable for humans to use.

In the branch of pattern recognition in speech processing, the author has proposed a framework of fuzzy logic system (FLS) mechanism that is applicable to specific speech pattern recognition subjects: speaker adaptation in speech recognition [3–7], speaker recognition [8], and acoustic event detection [9]. However, these developed FLS-based speech pattern recognition works lacked the characteristics of intelligence. The main weaknesses of such an FLS design is that a considerably prolonged time is required to train the FLS model, which will be of much difficulty in the convergence of FLS parameters especially when the FLS training data are scarce. An improved version of FLS, called "intelligent-FLS", has proposed to overcome these problems to promote the performance of the conventional FLS in intelligence [10]. Previous proposed intelligent-FLS contains two main components [10], the eigenspace-based FLS (Eigen-FLS) design for fast FLS model establishment with requiring only few training data from a speaker and a closed-loop FLS framework design with the learning ability of self-adjusting from inaccurate recognition results.

However, the Eigen-FLS design in previous developed intelligent-FLS [10] will still encounter an unreliable FLS model estimated when the collected training data for eigen value calculations of the FLS eigenspace (the eigen-decomposition process) is scarce. Speech pattern recognition applications with such the inaccurately calculated FLS model will not have a satisfactory recognition performance. To tackle this problem, this paper develops improved Eigen-FLS schemes by employing the popular linear interpolation technique, in which two methods are proposed, EigenMLLR-like Eigen-FLS and incremental MLED Eigen-FLS. The main difference between previous proposed Eigen-FLS in [10] and two improved Eigen-FLS in this work is that the model combination scheme of Eigen-FLS model representing a prior knowledge about training speakers and speaker specific model denoting the speaker-dependent knowledge about the testing speaker is introduced, and such the combination is performed by simple linear interpolation.

The linear interpolation technique [11] utilized in this work is particularly useful to deal with the problem of sparse training data for FLS model establishments since an inaccurately estimated Eigenspace-based FLS model due to insufficient training data could be immediately improved by interpolating directly certain degree of speaker-dependent FLS model. Such an interpolation by just simple linear combination is not expensive in calculation and therefore suitable for practical speech pattern recognition applications. Many approaches for speaker adaptation also employ such linear interpolation techniques, such as MAP of the Bayesian-based adaptation category [12] and MAPLR of the transformation-based adaptation category [13]. For improving the performance of those adaptation methods using linear interpolation, many studies have been proposed for the adaptation of Bayesian-based, transformation-based and eigenvoice-based categories. However, most of these studies rarely used the fuzzy mechanism. Fuzzy methods have been popular in lots of technical fields [14], including speech pattern recognition in this paper. Eigenspace-based MLLR (Eigen-MLLR) speaker adaptation is a typical representative of eigenvoice-based adaptation that uses the linear interpolation framework [15]. Following the idea of Eigen-MLLR, two proposed methods for enhancing previous developed Eigen-FLS, EigenMLLR-like Eigen-FLS and incremental MLED Eigen-FLS, use the similar formulation of linear interpolation where the fuzzy system is specifically used again to improve interpolations.

2. EIGENSPACE-BASED FUZZY LOGIC SYSTEM (EIGEN-FLS)

Eigenspace-based approaches, such as the famous eigenvoice speaker adaptation, and the typical image recognition techniques eigenface, eigenpalm, and eigenedge, have been frequently used in pattern recognition applications in recent years. The main purpose of using this category of eigenspace-based approaches is to provide an effective and fast learning scheme with few training data required for training recognition models. When model parameters are trained by the eigenspace-based method, the convergence problem of parameters is also resolved.

Ding's previous study [10] proposed an eigenspace-based method for speech pattern recognition applications using FLS, called "Eigen-FLS". The basic concept of the Eigen-FLS involves each number of speaker clusters being represented by a specific set of FLS parameters that can be developed in advance. The FLS model of the current target speaker is then represented as an interpolated form of the weighted sum of the speaker clusters. Such speaker-clustering-based FLS training could be considered speaker-space-based training. Mathematically, the estimated parameters of the sets of cluster models form the axes of speaker spaces, and by estimating an appropriate point for the speaker in the speaker space, the FLS model parameters for the speaker can be determined quickly. Under Eigen-FLS, *a priori* knowledge on the variations among all training speakers was represented as the set of FLS model parameters in the form of eigenvectors named "eigenfls". A new speaker model is then expressed as the linear combination of the set of eigenfls. Using the Eigen-FLS approach, the computing cost of an FLS establishment is reduced substantially, but remains capable of retaining the overall system characteristics to capture the variance between speakers.

The Eigen-FLS approach must ensure FLS eigenspace construction and coefficient estimation. In the FLS eigenspace construction phase (Fig. 1), N well-trained FLS models, each of which is derived from a speaker, must be established first. The model parameters of each FLS model are then "vectorized", forming a set of N "supervectors". Space dimension reduction techniques, such as principal components analysis (PCA), are then applied to the set of N supervectors to obtain N eigenvectors with dimension D, also called "eigenfls". Only the first K eigenfls are kept and are significant because they possess the most information from speech data, thus being capable of representing all considered variations. Finally, according to the K eigenfls, an accurate FLS space "K-space" is spanned and obtained. In the coefficient estimation phase, the FLS model of the target speaker is then acquired by estimating a set of weights to determine a weighted combination of eigenfls. In the eigen-decomposition work to determine the eigen value for each eigenfls, the maximum likelihood eigen-decomposition (MLED) algorithm is generally employed to perform the coefficient estimation due to its simplicity.

Let the supervector FLS^{EIGEN} of the new speaker be conducted in K-space as follows:

$$\mathrm{FLS}^{\mathrm{EIGEN}} = e(0) + w(1)\cdot e(1) + \cdots + w(K)\cdot e(K), \tag{1}$$

where $e(0)$ is the mean vector of N supervectors. The problem here is to estimate the weights $\{w(k), k = 1,2,\ldots,K\}$ that correspond to K eigenvectors $\{e(k), k = 1,2,\ldots,K\}$, to find a weighted combination of eigenfls. The MLED method used to derive the set of weight coefficients $\{w(k), k = 1,2\ldots,K\}$ using the speaker's specific training data X is to solve

$$\hat{w}_{\mathrm{ML}} = \arg\max P(X/w). \tag{2}$$

$\hat{w}_{\mathrm{ML}}$ in Eq. (2) can be solved by the expectation-maximization (E-M) algorithm.

However, the MLED algorithm would be ineffective in this case of scarce data from the test speaker. Enhancing Eigen-FLS to effectively tackle this problem is of great urgency, which would be introduced in the following section.

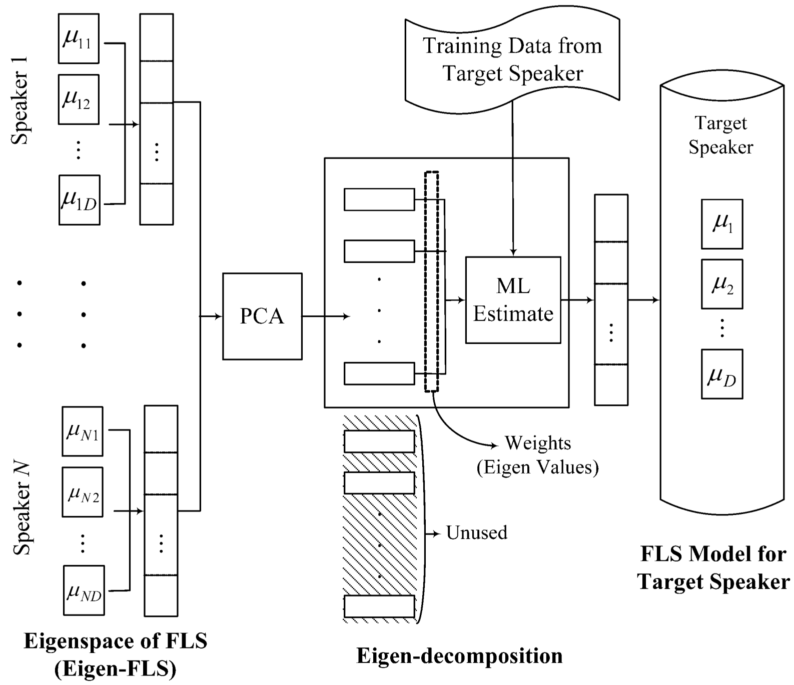

Fig. 1. Eigenspace-based FLS (Eigen-FLS) frameworks for speech pattern recognition.

3. ENHANCED EIGEN-FLS USING LINEAR INTERPOLATION FOR SPEECH PATTERN RECOGNITION

The Eigen-FLS model parameters determined using the improved Eigen-FLS proposed in this work will be much more accurate than those calculated by the conventional Eigen-FLS. This section presents two developed improved Eigen-FLS approaches, EigenMLLR-like Eigen-FLS and incremental MLED Eigen-FLS, both of which avoid the inaccurately determined Eigen-FLS model that is established using insufficient training data from the target speaker and contribute to achieve optimal recognition performance for FLS-based speaker adapted speech recognition, FLS-based speaker recognition and FLS-based acoustic event detection applications.

3.1. Eigenspace-based MLLR (Eigen-MLLR) Speaker Adaptation

Transformation-based model adaptation first derives appropriate transformations from a set of adaptation utterances acquired from a new speaker, and then applies them to clusters of HMM parameters. Leggetter et al. first proposed MLLR adaptation under the framework of affine transformation [16], and this method has become quite popular and successful due to its rapid adaptation. However, sufficient adaptation data is required to ensure accurate MLLR transformation estimation.

Eigenspaced-based MLLR is the eigenvoice version of MLLR, and involves a hybrid of MLLR-eigenvoice adaptation [15]. Speaker adaptive training (SAT) is a type of speaker clustering methods and is very similar to Eigenspace-based MLLR. However, these two approaches are apparently different. The goal of SAT is to reduce inter-speaker variability within the training set, and that of Eigenspace-based MLLR is to take advantage of prior knowledge about the test speaker's linear transforms. Doumpiotis and Deng suggested a combination of SAT and Eigenspace-based MLLR. The Eigen-MLLR adaptation framework essentially belongs to the type of the linear interpolation of a prior knowledge about training speakers and a speaker

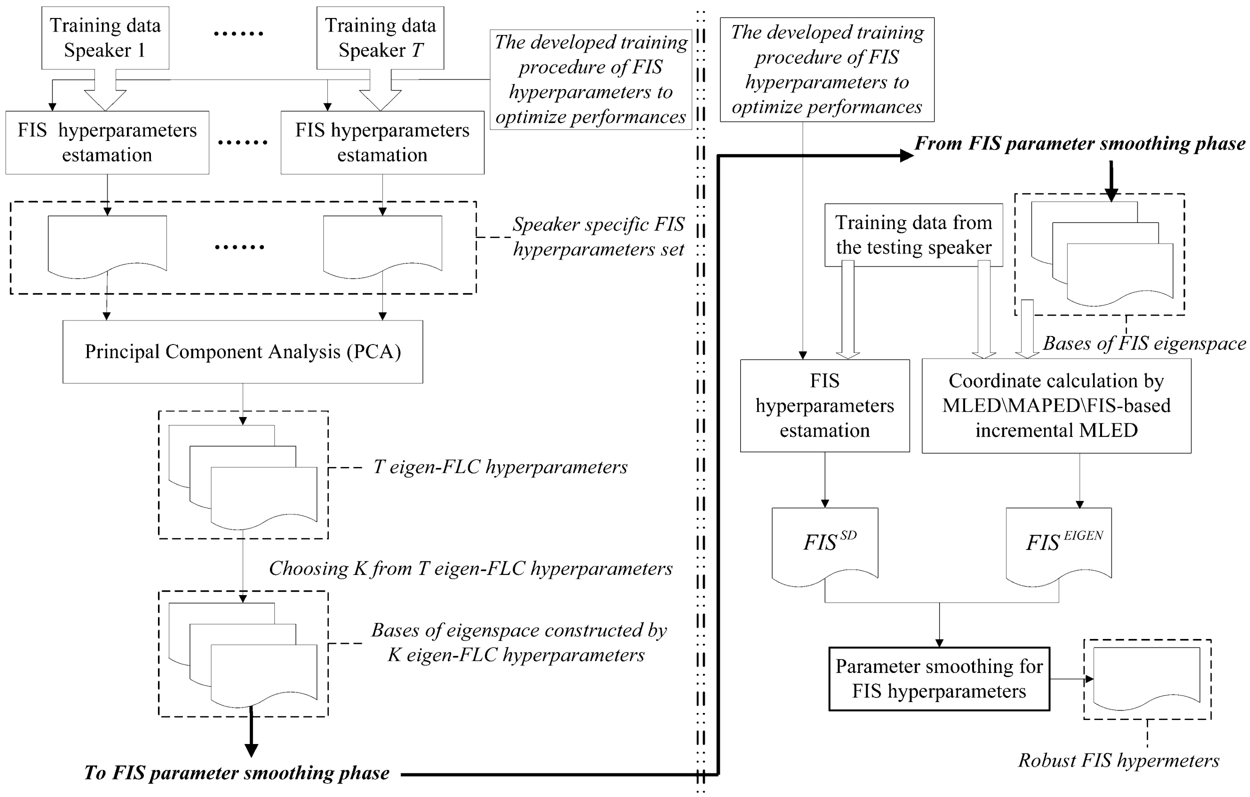

Fig. 2. The framework of the proposed EigenMLLR-like Eigen-FLS.

specific model as follows:

$$\tilde{W}_s = \frac{\tau \cdot W_s^{\mathrm{EIGEN}} + \sum_{t=1}^{N} \gamma_s(t) W_s^{\mathrm{MLLR}}}{\tau + \sum_{i=1}^{N} \gamma_s(t)}, \tag{3}$$

where a parameter smoothing procedure for the transformation matrices W_s^{EIGEN} and W_s^{MLLR} being carried out, and $\gamma_s(t)$ denotes the state occupation probability at time t for observation data $O = o_1, o_2, \ldots, o_t \ldots o_N$ and τ represents an empirically determined parameter

3.2. An EigenMLLR-like Approach for Eigen-FLS by Linear Interpolation

The EigenMLLR-like Eigen-FLS approach developed in this study to enhance the traditional Eigen-FLS is shown in Eq. (4), which is very similar to the one in Eq. (3), essentially an EigenMLLR-like estimation [11],

$$FLS = \alpha \cdot FLS^{\mathrm{EIGEN}} + (1-\alpha) \cdot FLS^{\mathrm{SD}}. \tag{4}$$

Figure 2 illustrates the framework of the proposed EigenMLLR-like Eigen-FLS method. Note that the fuzzy logic system FLS for a new speaker is essentially a weighted average of the *a priori* knowledge of the fuzzy logic system parameters, FLS^{EIGEN}, and the speaker dependent (SD) fuzzy logic system parameters, KLS^{SD}. FLS^{EIGEN} parameters are calculated only by the utterances of the target speaker without referencing any information from the speaker independent (SI) eigenfls space. As could be seen in Eq. (4), the linear combination parameter α controls the interpolation between the FLS^{EIGEN}-term and the FLS^{SD}-term. When a large number of training samples from the target speaker is available for FLS establishments, the value of α in Eq.uation (4) should be designed to be as small as possible so that the fuzzy logic system parameters FLS for the speaker could converge asymptotically to FLS^{SD}. Conversely, when training samples are extremely few (i.e., no sufficient training data are available for FLS training), the FLS establishment should be simply performed by referencing more prior fuzzy logic system parameters FLS^{EIGEN} (i.e., the α value in Eq. (4) should be given to approach 1). The robustness of the estimated Eigen-FLS model using Eq. (4) for a new speaker against a relatively small number of training data could then be ensured.

3.3. Improving the Accuracy of MLED-derived Eigen Values Using Linear Interpolation

As mentioned before, the coefficient estimation phase of Eigen-FLS performs a fast FLS establishment using the MLED eigen-decomposition algorithm to estimate a set of weights to find a weighted combination of eigenfls for the new speaker. Given sufficient training data, eigen-decomposition by the MLED method is effective. However, given insufficient training data, the accuracy of the estimated combination coefficients is dubious. Poor estimation of the combination coefficients in turn leads to incorrect positioning in the FLS eigen-space. The problem of scarce training data can be alleviated by using the MAPED scheme if heavy computation is permissible.

Given insufficient training data, it is necessary to be more "conservative" in using the combination coefficients thus derived. In other words, the effect of the improper data should be restricted so that the final FLS model does not reference too much from the combination coefficients derived with the insufficient training data. Therefore, this study proposes the following incremental MLED Eigen-FLS approach [11]

$$FLS = e(0) + \sum_{k=1}^{K} [\beta \cdot w(k) + (1_{\beta}) \cdot \mu_{k(w)}] e(k), \quad 0 \le \beta \le 1, \tag{5}$$

where $w(k)$ is the combination coefficient calculated by MLED and $\mu_{w(k)}$ is the prior mean of the combination coefficient. The linear combination cefficients for eigenvector decomposition are not calculated as in the maximum likelihood criterion. Instead, this approach calculates a weighted sum of the maximum likelihood estimate and the prior mean of the combination coefficient. The form of incremental MLED Eigen-FLS in Eq. (5) is very similar to the aforementioned EigenMLLR-like Eigen-FLS approach (see Eq. 4). A weight parameter β governs the balance of $w(k)$ and $\mu_{w(k)}$, mimicking the role of the parameter α in EigenMLLR-like Eigen-FLS. Using a weighting scheme with the adjustable parameter β should achieve satisfactory adaptation performance even when only a small amount of training data is available for eigen-decomposition. Note that the weight β varies depending on how much confidence one has in the combination coefficient derived from MLED. A possibly not so well estimate of the combination coefficient from MLED due to insufficient training data would preferably goes with β approaching 0 so that the biased estimate of $w(k)$ will be restricted. Conversely, 1-approaching β should be given.

3.4. Weighted Parameters α and β Regulated by an FLS

For the specific problem of speech pattern recognition using the developed EigenMLLR-like Eigen-FLS approach, the rules governing the parameter α regulation, given the input N, the number of the training samples, can be formulated as the following implications:

Rule 1: If N is small, then α is small,

Rule 2: If N is medium, then α is medium,

Rule 3: If N is large, then α is large.

Within the framework of the fuzzy process, the statements of linguistic terms with uncertainty to some degree can be formulated in quantized forms for subsequent computations. The formulation of the above implications is given as a set of three fuzzy IF-THEN fuzzy rules and the system output α.

Rule 1: If N isM_1, then $\alpha = f_1(N)$,

Rule 2: If N is $M_2(N)$, then $\alpha = f_2(N)$,

Rule 3: If N is $M_3(N)$, then $\alpha = f_3(N)$,

where $M_1(N), M_2(N)$, and $M_3(N)$ are the membership functions representing the degree of how much N is involved in the classes of linguistically "small", "medium" and "large", respectively, and are defined as

$$M_1(N) = \begin{cases} 1 & N \le N_1, \\ \frac{N_2 - N}{N_2 - N_1} & N_1 \le N \le N_2, M_2(N), \\ 0 & N \ge N_2, \end{cases}$$

$$M_2(N) = \begin{cases} 0 & N \le N_1 \text{or} N \ge N_3, \\ \frac{N - N_1}{N_2 - N_1} & N_1 < N \le N_2, \\ \frac{N_3 - N}{N_3 - N_2} & N_2 \le N < N_3, \end{cases}$$

$$M_3(N) = \begin{cases} 0 & N \le N_2, \\ \frac{N - N_2}{N_3 - N_2} & N_2 < N < N_3, \\ 1 & N \ge N_3. \end{cases} \tag{6}$$

$f_i(N)$, $i = 1, 2, 3$ are output functions in each rule for regulating the α value and are defined as

$$f_1(N) = a_1 \cdot N + b_1, \quad f_2(N) = a_2 \cdot N + b_2, \quad f_3(N) = a_3 \cdot N + b_3. \tag{7}$$

For the system output, α is defined as

$$\alpha = \frac{\sum_{i=1}^{3} M_i(N) f_i(N)}{\sum_{i=1}^{3} M_i(N)}. \tag{8}$$

The parameter β regulation in incremental MLED Eigen-FLS could also be carried out by an FLS as the parameter α is governed under a fuzzy process.

3.5. Case Study: Enhanced Eigen-FLS for MAP Speaker Adaptation

Speaker adaptation (SA) is a popular technique for speech recognition. The main process of speaker adaptation is to use sample utterances collected from a new speaker (the end-user of the system) for adapting the system internal parameter settings of the pre-established speech recognition model.

Maximum *a posteriori* (MAP) adaptation is a type of direct model adaptation that attempts to re-estimate the model parameters directly, and re-estimates only the portion of the model parameter units associated with the adaptation data [12]. The MAP estimate of the mean is a weighted average of the prior mean and sample mean, and the weights are functions of the number of adaptation samples if an adaptation speed parameter τ is fixed to yield the bias between the prior mean and the sample mean. To further enhance robustness of MAP adaptation, an FCMAP method was proposed in the author's previous study [4] to modify the value of τ, according to the amount of adaptation data. In FCMAP, a fuzzycontrol mechanism (a typical FLS) was developed to determine the appropriate τ for regulating the adaptation speed effectively.

In this work, the classical FLS of FCMAP is replaced with the fuzzy logic system determined by each of these two improved-versioned Eigen-FLS methods mentioned above, EigenMLLR-like Eigen-FLS and incremental MLED Eigen-FLS, for deriving a more accurate value and then increasing the effectiveness on speech model learning. These two improved MAP in this study name as EigenMLLR-like Eigen-FLSMAP and incremental MLED Eigen-FLSMAP.

4. EXPERIMENTS AND RESULTS

The proposed improved eigenspace-based fuzzy logic system methods by linear interpolation, EigenMLLR-like Eigen-FLS and incremental MLED Eigen-FLS, are performed on MAP speaker adaptation experiments for demonstrating their effectiveness.

4.1. System Description and Database

The speech signal was sampled at 8 kHz. The analysis frames were 30-ms wide with a 20-ms overlap. For each frame, a 24-dimensional feature vector was extracted. The feature vector for each frame was composed of a 12-dimensional mel-cepstral vector and a 12-dimensional delta-mel-cepstral vector.

The database MAT400 sub-database DB3 [17] was used to build up the initial speaker independent (SI) models in the form of a set of hidden Markov model (HMM) parameters, and a group of 10 speakers was summoned for utterance recording. 30 utterances of city names (one utterance for each of 30 cities) as adaptation data for setting up SA models were collected from each of the 10 speakers. In the recognition experiments, adaptation and testing data were gathered from a new group of five speakers. All the uttered data were recorded by a close-talking microphone.

The adaptation data consisted of 30 utterances from each speaker (one utterance for each of 30 cities). The testing data consisted of 60 utterances from the speakers, each uttering twice for 30 city names. For each speaker, 2, 6, 10, 14, 18, 22, 26, and 30 utterances were picked out from his/her 30-utterance adaptation data for SI model adaptation, and 8 sets of SA models are established per speaker. A total of 40 SA models are thus set up and used for performance comparison among conventional MAP, FCMAP, Eigen-FLS MAP, EigenMLLR-like Eigen-FLSMAP and incremental MLED Eigen-FLSMAP adaptation methods.

4.2. Experimental Results

In general, conventional MAP with $\tau = 30$ would have the best performance among various settings of τ and therefore, MAP with τ set as 30 will be chosen for performance comparison. Speaker adaptation experiments were carried out for each of the 5 speakers using conventional MAP with $\tau = 30$, FCMAP, Eigen-FLS MAP, proposed EigenMLLR-like Eigen-FLSMAP and developed incremental MLED Eigen-FLSMAP. Each of these adaptation methods is to adjust the SI speaker model to produce the associated SA models of eight, and the recognition performances are given in Fig. 3, of which each curve of adaptation methods shows the increasing condition of the averaged recognition rate and the learning speed of the speech model by the 5 testing speakers at all eight test cases.

As depicted in Fig. 3, recognition performances demonstrate the superiority of both the presented EigenMLLR-like Eigen-FLSMAP and incremental MLED Eigen-FLSMAP. It is clearly seen that the proposed EigenMLLR-like Eigen-FLSMAP as well as incremental MLED Eigen-FLSMAP has a better adaptive learning curve. For the conventional MAP, when the amount of training data is insufficient, the recognition rate is low, even lower than the baseline. For FCMAP and Eigen-FLS MAP, both the recognition rate curves are still a little dissatisfactory that the recognition accuracy maintains almost the same as the baseline when adaptation data is small and could not be competitive in the case of large adaptation data size. In contrast, the recognition rate of both the EigenMLLR-like Eigen-FLSMAP and incremental MLED Eigen-FLSMAP keeps better than the baseline when the amount of training data is not sufficient. Furthermore, when the amount of training data is increasing, the recognition performance of the conventional MAP, FCMAP and Eigen-FLS MAP become better than the baseline but still a little worse than that of EigenMLLR-like Eigen-FLSMAP and incremental MLED Eigen-FLSMAP.

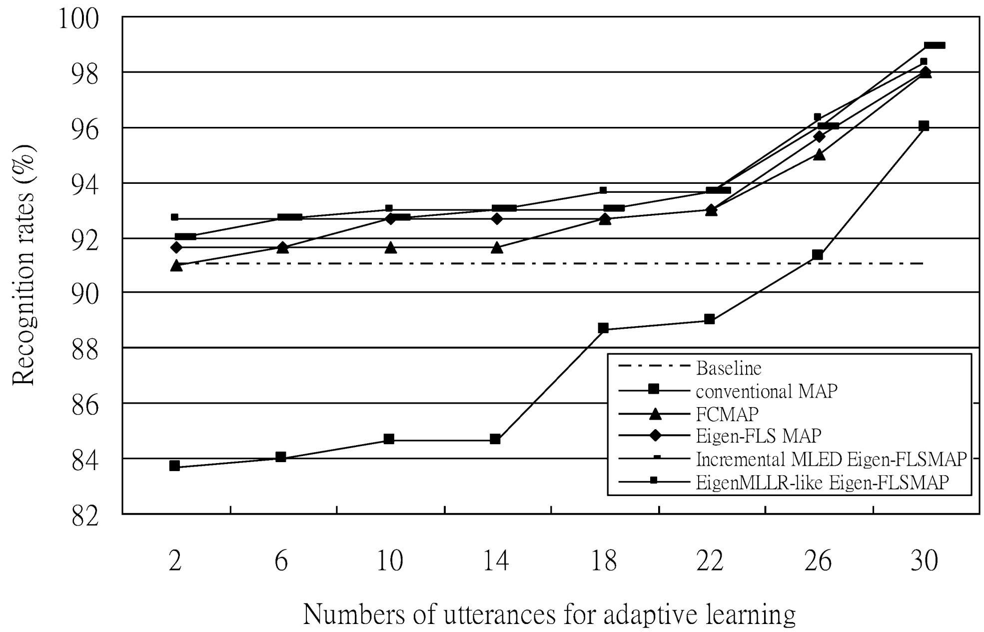

Fig. 3. Average recognition rate curves by five speakers using conventional MAP, FCMAP, Eigen-FLS MAP, EigenMLLR-like Eigen-FLSMAP and incremental MLED Eigen-FLSMAP adaptation methods.

5. CONCLUSIONS

This paper developed two approaches for enhancing the traditional eigenspace-based fuzzy logic system design, incremental MLED Eigen-FLS and EigenMLLR-like Eigen-FLS. Both of the proposed methods utilize a linear interpolation scheme with the linear combination coefficient effectively governed by an FLS. For a target speaker, the Eigen-FLS model parameters determined using the improved Eigen-FLS would be much more accurate than those calculated by the conventional Eigen-FLS. FLS-based speech pattern recognition applications use the developed incremental MLED Eigen-FLS and EigenMLLR-like Eigen-FLS would have more robustness against data scarcity from a test speaker and could then achieve optimal performances in recognition accuracy.

ACKNOWLEDGEMENTS

This research is partially supported by the National Science Council (NSC) in Taiwan under grant NSC 101-2221-E-150-084.

REFERENCES

1. Prasad, J.R., “Trends in handwriting recognition”, in *Proceedings of International Conference on Emerging Trends in Engineering and Technology*, pp. 491–495, 2010.
2. Gaikwad, S.K., Gawali, B.W., Yannawar, P., “A review on speech recognition technique”, *International Journal of Computer Applications*, Vol. 10, No. 3, pp. 16–24, 2010.
3. Ding, I.J., “Fuzzy logic based control for VFS speaker adaptation”, *International Journal of Fuzzy Systems*, Vol. 10, No. 4, pp. 292–297, 2008.
4. Juang, Y.T., Huang, K.C. and Ding, I.J., “Speaker adaptation based on MAP estimation using fuzzy controller”, *Pattern Recognition Letters*, Vol. 24, No. 15, pp. 2807–2813, 2003.
5. Ding, I.J., “Incremental MLLR speaker adaptation by fuzzy logic control”, *Pattern Recognition*, Vol. 40, No. 11, pp. 3110–3119, 2007.

6. Ding, I.J., “Enhancements of maximum likelihood eigen-decomposition using fuzzy logic control for eigenvoice-based speaker adaptation”,*International Journal of Innovative Computing, Information and Control*, Vol. 7, No. 7(B), pp. 4207–4222, 2011.
7. Ding, I.J., “Enhancing eigenspace-based MLLR speaker adaptation using a fuzzy logic learning control scheme”, *IEICE Transactions on Information and Systems*, Vol. E94-D, No. 10, pp. 1909–1916, 2011.
8. Ding, I.J., “Speaker verification using combined HMM-SVM under fuzzy system regulation”, *ICIC Express Letters*, Vol. 6, No. 2, pp. 557–562, 2012.
9. Ding, I.J., “Events detection for audio based surveillance by variable-sized decision windows using fuzzy logic control”, *Tamkang Journal of Science and Engineering*, Vol. 12, No. 3, pp. 299–308, 2009.
10. Ding, I.J., “Fuzzy logic system-based pattern recognition in speech processing: techniques and challenges”, *Information-An International Interdisciplinary Journal*, Vol. 25, No. 1, pp. 49–56, 2013.
11. Benediktsson, J.A. and Swain, P.H., “Consensus theoretic classification methods”, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, No. 4, pp. 688–704, 1992.
12. Lee, C.H., Lin, C.H. and Juang, B.H., “A study on speaker adaptation of the parameters of continuous density hidden Markov models”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 39, No. 4, pp. 806–814, 1991.
13. Chesta, C., Siohan, O., and Lee, C.H., “Maximum a posteriori linear regression for hidden Markov model adaptation”, *Proceedings of European Conference on Speech Communication and Technology*, pp. 211–214, 1999.
14. Soliman, M.A., Kaldas, M.M.S., Barton, D.C. and Brooks, P.C., “Fuzzy-skyhook control for active suspension systems applied to a full vehicle model”, *International Journal of Engineering and Technology Innovation*, Vol. 2, No. 2, pp. 85–96, 2012.
15. Chen, K.T., Liau, W.W., Wang, H.M., and Lee, L.S., “Fast speaker adaptation using eigenspace-based maximum likelihood linear regression”, in *Proceedings of International Conference on Spoken Language Processing*, pp. 742–745, 2000.
16. Leggetter, C.J. and Woodland, P.C., “Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models”, *Computer Speech and Language*, Vol. 9, pp. 171–185, 1995.
17. Wang, H.C., “MAT – A project to collect Mandarin speech data through telephone networks in Taiwan”, *Comput. Linguist. Chinese Lang. Process.*, Vol. 2, pp. 73–89, 1997.