Self Organization Mapping-Based Robust Speaker Identification in Noisy Environment using Linear Discriminant Analysis

Md. Rabiul Islam and Md. Fayzur Rahman
Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Bangladesh
Department of Electrical & Electronic Engineering, Daffodil International University, Bangladesh.
E-mail: rabiul_cse@yahoo.com and mfrahman3@yahoo.com

Abstract— The aim of this work is to evaluate the performance of the noise robust speaker identification system using Kohonen self organization mapping based neural network algorithm. Speaker identification system performs well under noiseless environment but its performance degrades when the environments become noisy. In this work, wiener filtering technique has used to reduce the white Gaussian environmental speech noises, standard feature extraction techniques have been applied to extract the effective speech features. Linear Discriminant Analysis based dimensionality reduction technique has been used to reduce the dimension of the extracted speech features. Kohonen self organization learning neural network based algorithm has applied for the learning and recognition model of the proposed system. Finally, NOIZEUS speech dataset has used to measure the efficiency of the proposed system under various environmental noisy conditions with different noise addition rate.

Keywords—noise robust speaker identification, self organization mapping, artificial neural networks, dimensionality reduction linear discriminant analysis.

Manuscript received on 17 March 2015 and accepted for publication on 9 June 2015.

1 INTRODUCTION

Speaker identification is a dynamic biometric task where both biological and behavioral biometric modalities are concatenated through speech feature extraction, learning and recognition modeling [1, 2]. Though biological and behavioral biometric modalities are combined through speaker identification task, it can achieve great efficiency than other individual biological and behavioral biometric system such as face, iris, retinal scan, fingerprint, hand geometry etc. Generally speaker identification system performs well under neutral conditions but its performance degrades when the environments become noisy. There are various techniques that have been developed by different researchers for automatic speech and/or speaker identification problem [3, 4, 5, 6, 7, 8, 9]. From the last several years researchers investigates and tries to enhance performance of the speaker identification in noisy environments [10, 11, 12, 13, 14, 15].

Although there are recent studies to handle reverberation and additive noise in feature [15] and model domain [13, 14] for speaker identification systems, the compensation techniques with respect to noise and reverberation for speaker identification systems are still an open question [16]. There are many speech enhancement algorithms proposed for robust automatic speech recognition, most of them relying on the assumption that the additive noise is a stationary process which is not always true for real-world applications [8].

Not only single classification techniques are used for noise robust speech and speaker identification but also hybrid classification techniques have also been applied. Genetically optimized Hidden Markov Model (HMM) based hybrid technique [17] has been used for noise robust speaker identification system. Both audio and visual features are used for noise robust speaker identification purpose. Face and speech features are used with the combination of different types of fusion methods with neutral [18] and noisy environment [19]. Leap reading feature with audio feature have also been used for speaker identification [20].

In this paper, we propose a Kohonen self organizing mapping based approach for noise robust speaker identification which is much more capable than other existing system. To measure the performance, the proposed system has been tested under eight different real world noisy conditions such as airport, babble, car, exhibition, restaurant, street, train and train station with four different noise addition rate i.e., 0dB, 5dB, 10dB and 15 dB.

2 STRATEGY OF THE PROPOSED SPEAKER IDENTIFICATION SYSTEM

The block diagram of the proposed speaker identification system is shown in figure 1. After acquisition of the speech signal, it is required to eliminate the noise from the speech signal. Then some speech pre-processing techniques have been applied to pre-process the speech before feature extraction. Various standard speech parameterization techniques are applied to effectively extract the speech features and finally select best of them. The dimensions of the extracted speech features are very large. As a result, dimension reduction technique is used to reduce the speech feature vector. The reduced feature vector is now feed to the Kohonen self organizing mapping based neural network algorithm for learning and testing. Different parameters of Kohonen self organizing mapping based network has been tested for learning and finally optimum parameters are chosen to find out the final result.

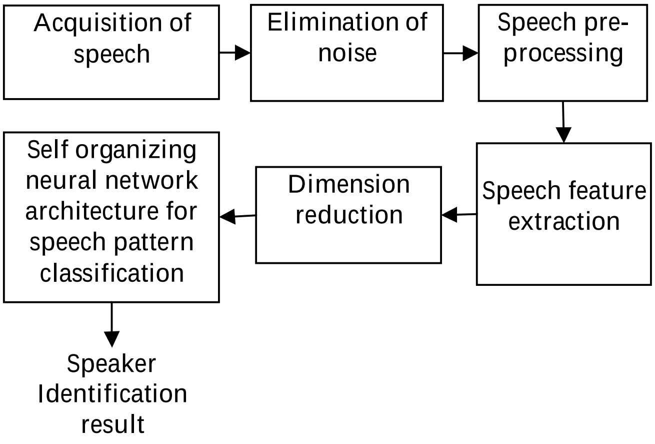

Figure 1: Architecture of the proposed speaker identification system.

3 Noise Elimination and Speech Pre-processing Techniques

Since speech signals generally captures a lot amount of environmental noises, it is necessary to reduce or eliminate noises from the captured speech signal. For this reason, wiener filtering technique has been used in the proposed system. The wiener filter is a noise removing filter based on Fourier iteration. Its main advantage is the short computational time it takes to find a solution [21].

Some pre-processing techniques are necessary to prepare the speech as an input for the system. Speech start and end points detection and silence part removal algorithm [22] have been used to extract the core part of the speech utterances. Pre-emphasis has been used to balance the spectrum of voiced sounds that have a steep roll-off in the high frequency region [23]. Since Sort Time Fourier Transform (STFT) is the appropriate method to analyze the speech features, windowing technique has been used to reduce the effect of the spectral artifacts that results from the framing process. Frame length of 10-30 milliseconds speech has been considered here and 25%-75% frame overlapping has been tested to get the optimum output of the speaker identification result.

4 Speech Feature Extraction and LDA based Dimension Reduction Technique

To extract the features from the speech utterances, various speech feature extraction techniques have been applied and find out the optimum results which will be focused in the experimental result section. Linear Prediction Coefficients (LPCs), Linear Prediction Cepstral Coefficients (LPCCs), Mel Frequency Cepstral Coefficients (MFCCs), Delta MFCCs and Delta Delta MFCCs [24] are tested to find out the best speech feature extraction technique for the proposed speaker identification system. The process of feature extraction and dimensionality reduction technique is shown in figure 2.

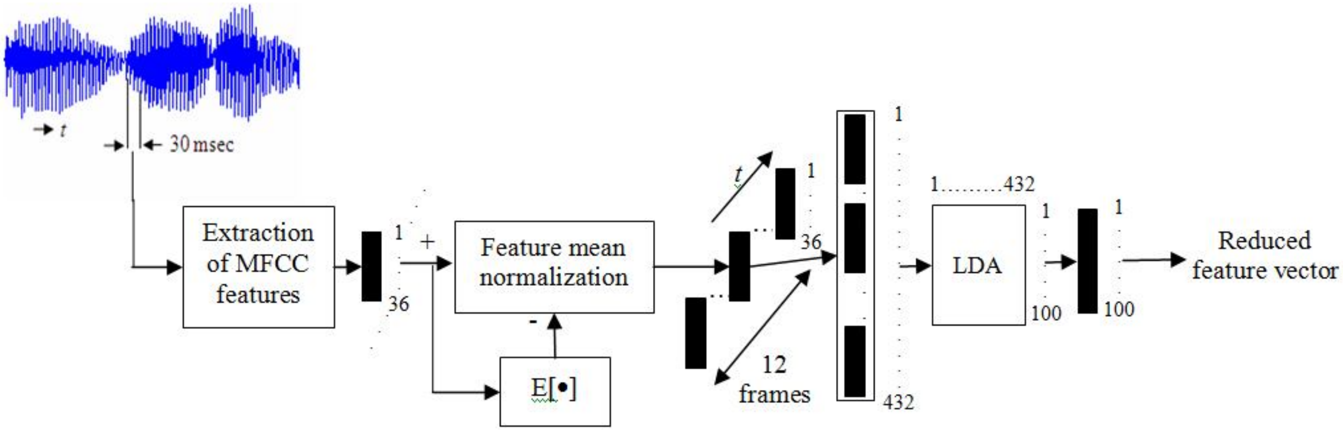

Figure 2: Feature extraction and dimensionality reduction from the speech utterance.

The dimensionality of the extracted speech feature vector is large. As a result, it is necessary to reduce the dimension of the speech features. To reduce the dimension of the speech feature vector, Linear Discriminant Analysis (LDA) based dimension reduction technique has been used.

The first step is to formulate the data sets and test sets, which are to be classified in the original space. The mean of each data set and mean of entire data set is to be calculated. After the mean image has been calculated, mean image is subtracted from each of those images of the training data set. Mean image is subtracted from each original image F_i and stored in the variable Φ_i. Each image in the data set differs from the average face by the vector Φ_i.

$$\Phi_i = F_1 - \Psi \tag{1}$$

Since LDA calculates the difference of features within all images of each person individually. So, one scatter matrix is calculated for each person from its images.

$$S_i = \sum_{j=1}^{L} F_{ij} F_{ij}^T \quad (2)$$

Here,

S_i is the scatter matrix of i^{th} person

L is the number of images of each person

F_{ij} is the j^{th} image of i^{th} person

Summation of all scatter matrices is called within-class scatter matrix which represents variation among images of each persons.

Within-class scatter matrix,

$$S_w = \sum_{i=1}^{M} S_i \quad (3)$$

Here,

M is the number of total persons

S_i is the i^{th} scatter matrix

Between-class scatter matrix represents the variation among persons. For Between-class scatter matrix,

$$S_B = 2\sum_{i=1}^{M} F_{mean,i} F_{mean,i}^T \quad (4)$$

Here,

M is the number of total persons

S_i is the i^{th} scatter matrix

$F_{mean,i}$ represents mean image of i^{th} person

Since LDA maximizes between-class scatter whereas minimizes the within-class scatter. To accomplish this, we must maximize W matrix where,

$$J(W) = |W^T S_B W| / |W^T S_w W| \quad (5)$$

From the matrix W, we will compute eigenvectors (Fishervectors) which will represents linear discriminant features of each person. The steps to compute eigenvectors from W matrix are given below:

1. Columns of W are eigenvectors satisfying the equation given below:

$$S_B W_i = \lambda_i S_W W_i \quad (6)$$

2. Eigenvalues are roots of the equation given below:

$$|S_B - \lambda_i S_W| = 0 \quad (7)$$

3. Calculation of eigenvectors by solving the equation given below:

$$(S_B - \lambda_i S_W) W_i = 0 \quad (8)$$

Eigenvectors of highest eigenvalues are selected and eigenvectors with lowest eigenvalues of the data set are ignored. Once eigenvectors are found, the next step is to order them by eigenvalue, highest to lowest. This gives the components in order of significance. Now those components having less eigenvalue can be ignored. If the eigenvalues are small, then it contains a less information about the data. To be precise, if original data have n dimensions in data set and so, n eigenvectors and eigenvalues are gained and then only the first p eigenvectors are chosen then the final data set has only p dimensions.

Now the feature vector is to be calculated. Taking the eigenvectors that we want to keep from the list of eigenvectors and forming a matrix with these eigenvectors in the columns construct this. At first eigenvectors are converted in column vector and then each of them are placed on a matrix in each row.

$$\text{Feature vector} = (eig_1\ eig_2\ eig_3 \ldots\ldots\ldots\ldots\ eig_p) \quad (9)$$

Finally, we get the feature vector in reduced dimension which can be used in classification process. Linear Discriminant Analysis (LDA) searches for those vectors in the underlying space that best discriminate among classes (rather than those that best describe the data). More formally, given a number of independent features relative to which the data is described, LDA creates a linear combination of these which yields the largest mean differences between the desired classes. Thus theoretically, LDA should give better performance than PCA [25, 26, 27].

5 SELF ORGANIZATION MAPPING BASED LEARNING AND TESTING MODEL

Kohonen self organizing mapping based learning and testing model has been developed for the proposed speaker identification system. The SOM is unlike most classification or clustering techniques in that it provides a topological ordering [28] of the classes. The architectural view of Kohonen self-organization map is shown on figure 3. The output grid contains the nodes where each of the node will win for each of the input set. For weight adaptation, the following equation has been used.

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t)(x_i(t) - w_{ij}(t)) \quad (10)$$

In this work, 200 nodes have been used on the output grid and the neighborhood size has been decreased according to the increment of adaptation process which is shown in figure 4. The learning procedure has been stopped when the minimum distance of a grid node is zero from the input nodes by using the equation,

$$d_j = \sum_{i=0}^{n-1} (x_i(t) - w_{ij}(t))^2 \quad (11)$$

Where d_j is the distance between the inputs and each output node j, x_i is the input to node i and w_{ij} the weight from input i to node j.

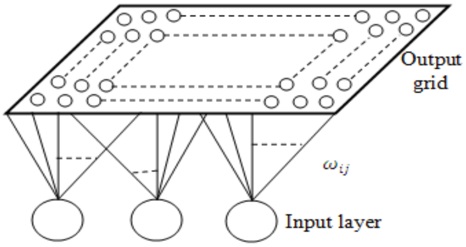

Figure 3: Architecture of Kohonen Self Organizing network.

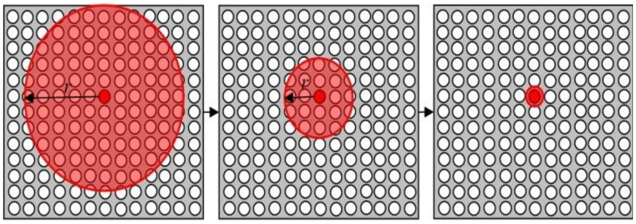

Figure 4: Reduction of topological neighborhood of Kohonen self-organizing network.

6 Experimental Results and Performance Analysis

NOIZEUS speech database has been used to measure the accuracy of the proposed speaker identification system. NOIZEUS speech corpus [29] was developed to facilitate comparison of speech enhancement algorithms among research groups. The noisy database contains 30 IEEE sentences (produced by three male and three female speakers) corrupted by eight different real-world noises at different SNRs. The noise was taken from the AURORA database and includes suburban train noise, babble, car, exhibition hall, restaurant, street, airport and train station noise.

Thirty sentences from the IEEE sentence database were recorded in a sound proof booth. The sentences were produced by three male and three female speakers. The IEEE database (720 sentences) was used as it contains phonetically balanced sentences with relatively low word context predictability. The thirty sentences were selected from the IEEE database so as to include all phonemes in the American English language. The sentences were originally sampled at 25 kHz and downsampled to 8 kHz [30].

One clean speech has been used for training and four others are used for testing purposes where noise addition varies from 0dB to 15dB with 5dB interval. The results of the proposed speaker identification system are shown in the table 1 to 9. The summarization of the results is shown on table 9. In this experiment, eight different real world environmental noises are added to successfully evaluate the performance of the proposed speaker identification system.

Table 1 to table 9 show the experimental results among various standard feature extraction techniques i.e., LPC, LPCC, MFCC, ΔMFCC, ΔΔMFCC. The results are taken with four different variations of SNRs such as 0dB, 5dB, 10dB and 15dB. To measure the overall performance of the proposed speaker identification, average performance is taken which is shown in table 9. From the average identification rate it has been found that MFCC can gives highest identification rate than other feature extraction method.

TABLE 1 Airport Noise Average Identification Rate (%) for NOIZEUS Speech Corpus

Method / SNR	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	87.00	88.00	64.75	68.00	69.25
10dB	81.25	85.50	57.50	62.25	65.00
5dB	67.75	75.00	48.25	54.50	58.25
0dB	59.00	67.25	50.00	55.50	50.50
Average	73.75	78.94	55.13	60.06	60.75

TABLE 2 Babble Noise Average Identification Rate (%) for NOIZEUS Speech Corpus

Method / SNR	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	75.00	87.50	57.75	56.25	72.00
10dB	71.25	83.50	47.25	48.00	68.25
5dB	57.50	67.25	36.50	47.75	62.25
0dB	55.00	55.50	36.67	44.25	57.25
Average	64.69	73.44	44.54	49.06	64.94

TABLE 3 Car Noise Average Identification Rate (%) for NOIZEUS Speech Corpus

Method / SNR	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	68.25	88.25	56.50	67.75	68.50
10dB	65.50	74.50	45.00	55.25	60.00
5dB	55.50	67.25	45.25	55.50	60.25
0dB	55.5-	58.00	39.00	45.00	55.00
Average	61.19	72.00	46.44	55.88	60.94

TABLE 4 Exhibition Hall Noise Average Identification Rate (%) for NOIZEUS Speech Corpus

Method / SNR	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	73.00	82.25	60.50	65.00	72.25
10dB	70.25	70.00	54.50	63.50	65.00
5dB	65.50	68.25	50.25	62.50	61.50
0dB	57.75	62.25	45.00	53.75	55.50
Average	66.63	70.69	52.56	61.19	63.56

TABLE 5 Restaurant Noise Average Identification Rate (%) for NOIZEUS Speech Corpus

Method / SNR	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	73.50	80.50	44.25	71.25	72.25
10dB	70.00	70.75	42.25	65.50	61.50
5dB	62.50	65.50	42.25	52.25	60.00
0dB	48.00	52.25	37.75	50.50	53.50
Average	63.50	67.25	41.63	59.88	61.81

TABLE 6 STREET NOISE AVERAGE IDENTIFICATION RATE (%) FOR NOIZEUS SPEECH CORPUS

SNR \ Method	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	75.50	80.25	52.25	67.75	75.00
10dB	67.75	70.00	45.50	53.50	65.25
5dB	65.50	69.25	42.25	67.75	63.25
0dB	53.25	65.50	35.50	55.50	52.25
Average	65.50	71.25	43.88	61.13	63.94

TABLE 7 TRAIN NOISE AVERAGE IDENTIFICATION RATE (%) FOR NOIZEUS SPEECH CORPUS

SNR \ Method	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	73.50	75.50	48.00	57.25	70.00
10dB	65.50	70.00	42.25	53.75	62.25
5dB	50.50	60.25	40.25	50.50	50.50
0dB	45.50	60.50	35.00	48.00	50.00
Average	58.75	66.56	41.38	52.38	58.19

TABLE 8 TRAIN STATION NOISE AVERAGE IDENTIFICATION RATE (%) FOR NOIZEUS SPEECH CORPUS

SNR \ Method	MFCC	ΔMFCC	ΔΔMFCC	LPC	LPCC
15dB	65.00	78.50	42.50	55.50	62.50
10dB	60.50	60.25	40.00	50.00	60.00
5dB	49.50	50.50	37.75	43.25	50.50
0dB	45.50	50.00	35.00	38.50	45.25
Average	55.13	59.81	38.81	46.81	54.56

TABLE 9 OVERALL AVERAGE SPEAKER IDENTIFICATION RATE (%) FOR NOIZEUS SPEECH CORPUS

Various Noises \ Method	MFCC	Δ MFCC	ΔΔ MFCC	LPC	LPCC
Airport Noise	73.75	66.57	64.60	62.78	61.45
Babble Noise	64.69	64.00	63.13	61.63	60.85
Car Noise	61.19	64.21	62.32	60.95	60.64
Exhibition Hall Noise	66.63	63.60	61.06	60.44	60.63
Restaurant Noise	63.50	60.72	60.01	60.39	60.83
Street Noise	65.50	60.76	60.41	60.78	61.01
Train Noise	58.75	59.17	60.29	60.91	61.09
Train Station Noise	55.13	59.39	60.85	61.22	61.18
Average Identification Rate (%)	63.64	62.30	61.58	61.14	60.96

7 CONCLUSIONS AND OBSERVATIONS

From the experimental results and performance analysis, it can say that the performance of the proposed self-organized mapping based speaker identification system is well enough for the real life applications. In this work, Linear Discriminant Analysis based dimensionality reduction technique has been used which can work well for inter-class variations. Effective noise removing technique and speech pre-processing techniques are used to achieve the highest efficiency. Though the highest identification rate of the proposed system has been found at 63.64%, the performance can be tested by using large speech database. Another feature extraction technique such as wavelet based speech feature extraction technique and other classification techniques i.e., Hidden Markov Model, Gaussian Mixture Model, Genetic Algorithm etc may be used to enhance the efficiency of the system.

REFERENCES

[1] S. Furui, "40 years of progress in automatic speaker recognition," *Lecture Notes Comput. Sci.,* vol. 5558, pp. 1050–1059, 2009.

[2] J. P. Campbell, "Speaker recognition: A tutorial," *Proc. IEEE*, vol. 85, no. 9, pp. 1437–1462, Sep. 1997.

[3] Jacobsen, J. D., "Probabilistic Speech Detection," *Informatics and Mathematical Modeling,* DTU, 2003.

[4] Jain, A., R.P.W.Duin, and J.Mao., "Statistical pattern recognition: a review," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 22, pp. 4–37, 2000.

[5] Reynolds, D.A., "Experimental evaluation of features for robust speaker identification," *IEEE Transactions on SAP,* vol. 2, pp. 639-643, 1994.

[6] Wu, D., Morris, A.C. & Koreman, J., "MLP Internal Representation as Disciminant Features for Improved Speaker Recognition," *Proc. NOLISP2005,* Barcelona, Spain, pp. 25-33, 2005.

[7] Picone, J., "Signal modeling techniques in speech recognition," *Proceedings of the IEEE,* vol. 81, no. 9, pp. 1215–1247, 1993.

[8] L.P. Cordella, P. Foggia, C. Sansone, M. Vento., "A Real-Time Text-Independent Speaker Identification System," *Proceedings of 12th International Conference on Image Analysis and Processing,* IEEE Computer Society Press, Mantova, Italy, pp. 632 – 637, 2003.

[9] D. Kewley-Port and Y. Zheng, "Auditory models of formant frequency discrimination for isolated vowels," *Journal of the Acostical Society of America,* vol. 103, no. 3, pp. 1654–1666, 1998.

[10] A. Solomonoff, W.M. Campbell, and I. Boardman, "Advances in channel compensation for SVM speaker recognition," *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing,* pp. 629–632, Philadelphia, USA, March, 2005.

[11] P. Kenny, P. Ouellet, N. Dehak, V. Gupta, and P. Dumouchel, "A study of interspeaker variability in speaker verification," *IEEE Trans. Audio, Speech and Language Processing,* vol. 16, no. 5, pp.980 –988, July 2008.

[12] N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification," *IEEE Trans. Audio, Speech and Language Processing,* vol. 19, no. 4, pp.788 –798, May 2011.

[13] J. Ming, T.J. Hazen, J.R. Glass, and D.A. Reynolds, "Robust speaker recognition in noisy conditions," *IEEE Trans. Audio, Speech and Language Processing,* vol. 15, no. 5, pp.1711 –1723, July 2007.

[14] D. Garcia-Romero, X. Zhou, and C. Y. Espy-Wilson, "Multicondition training of gaussian plda models in i-vector space for noise and reverberation robust speaker

recognition," *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing,* pp. 4157-4260, March 2012.

[15] C.M. Vannicola, B.Y. Smolenski, B. Battles, and P.A. Ardis, "Mitigation of reverberation on speaker identification via homomorphic filtering of the linear prediction residual," *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing,,* pp. 5512–5515, May 2011.

[16] R. Saeidi, A. Hurmalainen, T. Virtanen and D. A. van Leeuwen, "Exemplar-based Sparse Representation and Sparse Discrimination for Noise Robust Speaker Identification," *Proc. Odyssey: the speaker and language recognition workshop,* Singapore, June 2012.

[17] Md. Rabiul Islam, Md. Fayzur Rahman and M. Abdus Sobhan, "Genetically Optimized HMM for Robust Speaker Identification in Noise," *International Conference on Advances in Electrical Engineering,* Independent University, Bangladesh, pp. 258-263, Dec 19-20, 2011, Dhaka, Bangladesh.

[18] Gerasimos Potamianos, Chalapathy Neti, and Sabine Deligne, "Joint Audio-Visual Speech Processing for Recognition and Enhancement," *Auditory-Visual Speech Processing Tutorial and Research Workshop (AVSP),* pp. 95-104, St. Jorioz, France, Sep 2003.

[19] Md. Rabiul Islam and Md. Fayzur Rahman, "Hybrid Feature and Decision Fusion based Audio-Visual Speaker Identification in Challenging Environment," *International Journal of Computer Applications (IJCA)*, vol. 9, no. 5, pp. 9-15, Nov 2010.

[20] A. Adjoudant, C. Benoit, "On the integratio of auditory and visual parameters in an HMM-based ASR," *Speechreading by Humans and Machines: Models, Systems, and Speech Recognition, Technologies and Applications*, ed. D.G. Strok and M. E. Hennecke, (Springer, Berlin, Germany, 1996), pp. 461-472.

[21] Rosalind Wang, "Wiener Filtering," *PHYS 3301, Scientific Computing,* Project report for NOISE group, May 2000.

[22] Kitayama, K., Goto, M., Itou, K., and Kobayashi, T., "Speech Starter: Noise-Robust Endpoint Detection by Using Filled Pauses," *Eurospeech 2003*, Geneva, pp. 1237-1240, 2003.

[23] Qi, L., Zheng, J., Tsai, A., and Zhou, Q., "Robust Endpoint Detection and Energy Normalization for Real-Time Speech and Speaker Recognition", *IEEE Transaction on speech and Audion Processing*, vol. 10, no. 3, 2002.

[24] Davis, S. B., and Mermelstein, P., "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-28, no. 4, pp. 357–366, 1980.

[25] Peter N. belhumeur, Joao P. Hespanha, David J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transaction on Pattern Analysis and Machine Intelligence,* vol. 9, no. 7, pp. 711-720, July 1997.

[26] Kamran Etemad and Rama Chellappa, "Discriminant analysis for recognition of human face images," Department of Electrical Engineering and Center for Automation Research, University of Maryland, College Park, Maryland 20742, A/Vol. 14, No. 8, 1997.

[27] Hongjun Jia and Aleix M. Martinez, "Support Vector Machines in Face Recognition with Occlusions," Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210, USA.

[28] Teuvo Kohonen, "Self-Organizing formation of topologically correct feature maps," *Biol. Cybern.,* vol. 43, pp.59- 69, 1982.

[29] Hu, Y. and Loizou, P., "Subjective comparison of speech enhancement algorithms," *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP),* pp. 153-156, Toulouse, France, May 2006.

[30] Hu, Y. and Loizou, P., "Evaluation of objective quality measures for speech enhancement," *IEEE Transactions on Speech and Audio Processing,* vol. 16, no. 1, pp. 229-238, Jan 2008.

Md. Rabiul Islam was born in Rajshahi, Bangladesh, on December 26, 1981. He received his B.Sc. degree in Computer Science & Engineering and M.Sc. degrees in Electrical & Electronic Engineering in 2004, 2008, respectively from the Rajshahi University of Engineering & Technology, Bangladesh. From 2005 to till now, he is a faculty member in the Department of Computer Science & Engineering at Rajshahi University of Engineering & Technology. He is serving as an associate professor at the same place. He completed his Ph.D. degree programme under the excellent guidance of Dr. Md. Fayzur Rahman, Department of EEE at Rajshahi University of Engineering & Technology, Bangladesh. His research interests include bio-informatics, human-computer interaction, speaker identification and authentication under the neutral and noisy environments.

Md. Fayzur Rahman was born in 1960 in Thakurgaon, Bangladesh. He received the B. Sc. Engineering degree in Electrical & Electronic Engineering from Rajshahi Engineering College, Bangladesh in 1984 and M. Tech degree in Industrial Electronics from S. J. College of Engineering, Mysore, India in 1992. He received the Ph. D. degree in energy and environment electromagnetic from Yeungnam University, South Korea, in 2000. Following his graduation he joined again in his previous job in BIT Rajshahi. Currently he is a Professor in Electrical and Electronics Engineering in Daffodil International University, Bangladesh. He is currently engaged in education in the area of Electronics & Machine Control and Digital signal processing.