Missing Reliability Correction in Modality Information Integration for Robust Speaker Identification

M. Tariquzzaman[1], Kim Jin Young[1]*, Seung You Na[1], M. A. Rashid[2] and S. Tamaki[2],
[1]Chonnam National University, School of Electronics & Computer Engineering, Gwangju 500-757 Republic of Korea
[2]Dept. of Information Engineering, University of Ryukyus, Nishihara, Okinawa, 903-0213, Japan
* Corresponding Author

Summary
In the emerging biometrics technology, speaker identification in real environment is one of the key issues for enhancing the density of human computer interaction. In this paper, we propose an optimizing factor through a fuzzy membership function for correcting the reliability in different modalities reliability measure in a bimodal fusion process for speaker identification. In the bimodal speaker identification system, we have also applied our proposed extended modified convection reliability function to account optimal reliability simultaneously for audio and visual information integration. For creating mismatch in between train and test stage, babble noises and artificial illumination have been added to test speeches and lip images, respectively. Local PCA have been applied at features level at both classifiers system for reducing the dimension of features vector at different level of signal distortion. We have applied particle swarm optimization (PSO) for optimizing the proposed fuzzy membership function based optimizing factor and modified convection function's optimizing parameters. The speaker identification experiments have been performed using VidTimit database. Experimental results show that our proposed method enhanced the identification accuracy in comparison with the baseline system thus demonstrated the validation of the proposed approach and most notably maintains the consistency of the integration process.

Key words:
Speaker identification; fuzzy membership function; optimal reliability; local PCA; particle swarm optimization.

1. Introduction

In the 21st century, the need of biometrics appliance is continuously growing high for daily dealings and communication at different organizations such as government, public or private entities. Multimodal speaker identification, a goal of biometrics technology, are deploying increasingly in the secured organizations. Speaker identification is a task of identifying speaker using speaker behavioral/physiological traits such as speech signal and visual signal which are carrying linguistic/non-linguistic information and visual information respectively. The development of the speaker recognition technology particularly multimodal speaker recognition is still an active area of research which has a range of applications from national security, communication system security, computer security, computer network security, e-commerce to forensic [1-3]. Besides, increasing demands of intelligence interface with system like humanoid robot have created a need for automatic speaker recognition. With the current state-of-the-art biometrics technology, speaker identification may perform well under controlled environment. However, in real environment voice signal can easily be exposed of different degree of noises due to the channel distortion, its surrounding environments and codec distortion, meanwhile, the visual signal also can be degraded due to image quality, lighting condition and occlusion. As a result, the performance of unimodal speaker identification as well as the multimodal speaker identification system is degraded. Nonetheless, multimodal biometrics system performance is better than unimodal biometrics system with a high degree of careful design. For designing a multimodal biometrics system, a number of issues are focused which are mainly categorized into three; which modality to be fused, where to be fused and how the different modalities information to be fused including its reliability. Besides the single modality based speaker identification/verification system [4-8], there are different approach existing in the literature [9-12] for multimodal based speaker identification system where reliability/confidence measure and mismatch in train and test data are the important issues in fusion process. However, mostly in the existing integrated biometrics system, a high degree of mismatch was considered in the audio signal classifier system only.

On the other hand, local PCA has unique properties to reduce the dimension; as well has shown some robustness through mimicking the noisy conditions in speech and image signal in single modality based recognition system. However, PCA has no role to enhance the identification accuracy in different modality-based integrated scheme using reliability.

Manuscript received March 5, 2010
Manuscript revised March 20, 2010

Nevertheless, a high level of lighting condition and its directional properties could mislead the score-based reliability measure and thus affect the structure of integrated biometrics system and its performance. Obviously, the ultimate goal to measure the reliability is to reduce classification error rate and thus to increase the performance and robustness of a multimodal speaker identification system through maintaining the integration structure of multimodal biometrics system and reliability statistics.

In this paper, we will propose an optimizing factor in reliability measure through a fuzzy membership function to account misclassification due to observation equal likelihood probability simultaneously through the audio and lip information based reliability measure in bimodal speaker identification system. Moreover, we will apply the [18] modified convection reliability function's optimizing parameters to account optimal reliability simultaneously through the audio and lip information based reliability measure in bimodal speaker identification system. Babble noises and artificial illumination are added to testing speech and visual signal, respectively. For individual experts' speaker model generation, we have implemented the classical Gaussian mixture model (GMM) training [4]. The entire experiments are performed using VidTimit DB [13].

2. Classifier System Architecture

In this section, we will briefly describe the individual classifiers system and the overall baseline system.

2.1 Baseline System

A schematic diagram of an audio-visual integrated speaker identification system is shown in figure 1.

Certainly, in a bimodal speaker identification system there are two individual classifiers which are termed as audio based classifier/expert system and visual signal based classifier/expert system. Briefly the overall process can be described as follows:

The voice and visual image features are fed to its respective expert is termed as classifier. These classifiers calculate the set of likelihood score. The expert reliability parameter is estimated from each expert observation likelihood score. The fusion weighting parameters are calculated using the reliability parameters of different classifiers for integration. Thereafter, each classifier weight factor which is calculated differently from the likelihood score of each expert are weighted to each classifier score during integration and finally combined set of score is achieved. Finally, the highest score of likelihood is decided as the identified speaker from the combined score.

In the entire experiments, mel-frequency cepstrum coefficient (MFCC) and its delta features are extracted from speech signal and cepstral mean subtraction (CMS) [15] is performed on these features vector. Finally the local principal component analysis (PCA) [17] is applied on these robust features to obtain the reduced dimension features vector which are used in audio based classifier system. Similarly, discrete cosine transform (DCT) features of visual signal are also reduced by local PCA and which are used in the lip based classifier system. Manually we have taken the image region-of-interest (ROI) considering the lip center as the base point for ROI determination and thus created the lip database.

In individual classification process, each modality expert generates the log-likelihood score generally termed as $S(O_m | \lambda_n)$ where O is either voice or lip information. When m= 1, it indicates that it is audio information while

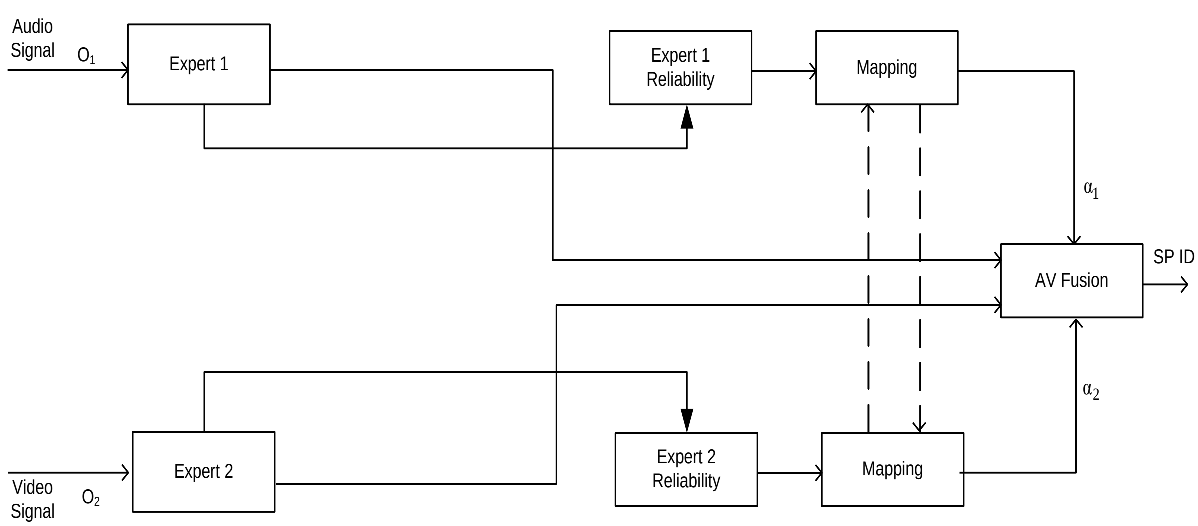

Figure 1. A schematic diagram of an audio-visual integrated system

at m=2 it is lip information and n of λ_n is the n-th speaker of the Gaussian Mixture Model (GMM) model and n=1,..., N.

2.2 Reliability measure and audio-visual information fusion

Broadly, the reliability measure uses the statistical nature of the experts which can be categorized into two i.e., the reliability parameters can be measured at either the signal level or at the expert score level. In literature, there exists some method for audio signal confidence measure and score-based both modalities reliability measure. However, there is a need to devise the missing reliability and its optimality in fusion process. The score-based reliability measures are mainly categorized: score entropy, dispersion, variance, cross classifier coherence and score difference. The score difference [11-12] is our special interest which has been taken as the baseline method in this paper. The baseline method takes the statistical nature of observation probability based on the rank of information. The entire process for reliability parameters calculation is followed by the following sequences:

Firstly, the individual modality generates the likelihood scores which are normalized using min-max normalization method differently. Mathematically we can express

$$\bar{S}(O_m \mid \lambda_j) = \frac{S(O_m \mid \lambda_j) - \mathrm{MinP}_m}{\mathrm{MaxP}_m - \mathrm{MinP}_m} \qquad (1)$$

where $\mathrm{MaxP}_m = \max_j S(O_m \mid \lambda_j)$ and $\mathrm{MinP}_m = \min_j S(O_m \mid \lambda_j)$

Secondly, the different modalities reliability are calculated using the highest rank of the normalized score, i.e.,

$$\rho_m = \overline{\mathrm{MaxP}}_m - \overline{\mathrm{Max2P}}_m \qquad (2)$$

where, $\overline{\mathrm{MaxP}}_m$ is the highest rank of the expert score; m represent the modality either audio or video and $\overline{\mathrm{Max2P}}_m$ is the second highest rank of the individual expert score in normalized form. After normalization $\overline{\mathrm{MaxP}}_m$ becomes 1 if there is no preferred range and the above equation can be written as

$$\rho_m = 1 - \overline{\mathrm{Max2P}}_m \qquad (3)$$

Thirdly, for mapping in between the reliability and the expert weighting factor, the weighting factor is calculated from the reliability function as follows:

$$\alpha_1 = \frac{\rho_1}{\rho_1 + \rho_2}, \quad \alpha_2 = 1 - \alpha_1 \qquad (4)$$

Fourthly, each modality score is integrated using the weighting factor as

$$S(O_1, O_2 \mid \lambda_n) = \alpha_1 \tilde{S}(O_1 \mid \lambda_n) + \alpha_2 \tilde{S}(O_2 \mid \lambda_n) \qquad (5)$$

Then finally the arg max is applied on the combined score to identify speaker.

3. Proposed Fuzzification in reliability and its optimization

3.1 Background and Introducing Optimizing Factor in Reliability Measure

From the above described section 2, we see that integrated score of observation are calculated using equations (1) to (5). Specifically, the reliability value which is expressed by equation (3) is determined from the audio and visual modality information individually and it has an important role to enhance the speaker recognition performance. For each expert, reliability value expressed by the equation (3) which was derived from the normalized observation probabilities and then the argument logic applied for speaker identification. Explicitly, even though the fusion final goal is to maximize the recognition rate; nonetheless, mathematically while, $\overline{\mathrm{MAXP}} = \overline{\mathrm{MAX2P}}$ ultimately equation (3) output will be 0. We can think two different possible conditions while $\overline{\mathrm{MAXP}} = \overline{\mathrm{MAX2P}}$ in the fusion process through the equation (3) to equation (5).

Case 1: The combined score expressed by equation (5) will be either audio signal based expert score or visual signal expert score which is not the desire of fusion process.

Case 2: if $\overline{\mathrm{MAXP}} = \overline{\mathrm{MAX2P}}$ arises on both case reliability measure at the same state then none of the expert score will be exist, as a results, misclassification will be certainly happened except the first identification which is practically based on its position rather than its observation state and this is also not the desire of multimodal biometrics technology .

Thus, we need to modify the reliability function so that the above two condition could be minimized and the ultimate goal of fusion to be fulfilled.

Therefore, we propose a fuzzy membership function based on $\overline{\mathrm{MAX2P}}$ which is as follows:

$$M(\tau)=\begin{cases} f & \text{if } \overline{\mathrm{MAXP}}=\overline{\mathrm{MAX2P}} \\ \overline{\mathrm{MAX2P}} & \text{otherwise}\end{cases} \qquad (6)$$

In equation (6), $M(\tau)$ is the fuzzy membership function and f is the optimized threshold. We can directly use the output of the fuzzy membership function in equation (3) instead of $\overline{\max 2P_m}$. To search the value of f in linear way is not appropriate thus we need some optimization.

3.2 Background and proposed extension of modified reliability function

From the above section II, we also see that the reliability value which is expressed by equation (3) is determined from the audio and visual modality information individually and it has an important role to enhance the speaker recognition performance. Explicitly, even though the convection function's i.e., $\rho_m = f(\tilde{S}(O_m \mid \lambda_1), \tilde{S}(O_m \mid \lambda_2), \ldots, \tilde{S}(O_m \mid \lambda_n))$ final goal is recognition rate; nonetheless, there is no optimization parameter that maximizes the identification rate (IR) in the bimodal speaker identification system. We can think two different possible conditions in the reliability measure regarding reliability function expressed by the equation (3).

a. Overestimation: Reliability value is estimated in higher order rather than its optimum level; In this case, we should regulate the reliability function so that the reliability value reaches at the optimum point.

b. Underestimation: Reliability value is estimated poorly from its ground truth; In this case, also we should regulate the reliability function to raise the reliability value. Introducing an optimization factor in reliability function, we can control the reliability function so that an improvement in audio-visual speaker identification could be achieved thus the ultimate goal of fusion to be fulfilled.

Thus, we have extended the proposed modified convection function in [18] through introducing the optimization factor on both modality i.e., introducing the optimization factors let f_m on $\overline{\max 2P_m}$, so that the above two case can be controlled. Mathematically, we can express the extended modified reliability function as follows:

$$\rho'_m = 1 - (\overline{\mathrm{Max}2P_m})^{f_m} \qquad (7)$$

In equation (7) f_m i.e., f_1 and f_2 are the optimization variables. Considering $\overline{\max 2P}$ we can conclude the following condition for the modified reliability values for different limit value of f_m.

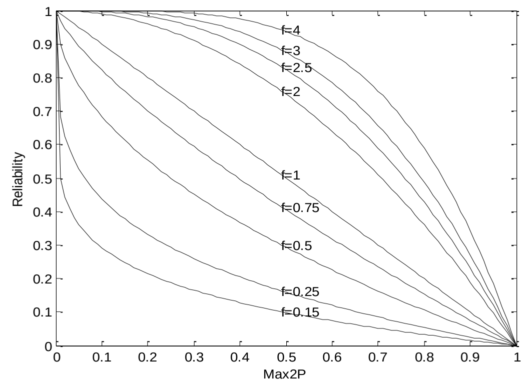

Figure 2. Graphical views of reliability values at Different f on $\overline{\mathrm{MAX2P}}$

- $0 \le f_m \le 1: \ \rho'_m < \rho_m$
- $f_m = 1: \rho'_m = \rho_m$
- $f_m > 1: \rho'_m > \rho_m$

Figure 2 shows the relationship and physical meaning between $\overline{\max 2P}$ and reliability values for different numerical values of the optimization factor f_1 or f_2.

However, we cannot get the optimum values f_m in linear searching way thus we need optimization. For optimization f, f_1 and f_2 we need a target function which is in detail in the following section.

3.3 Optimization target function

The optimization object function can be defined with the help of Speaker identification rate. The optimization object function is defined in the following:

$$x(f, f_1, f_2) = \frac{\sum_{k=1}^{K}\sum_{l=1}^{L_k} \delta(\arg \mathrm{Max}_j(P_j(X_{kl})), k)}{K\sum_{K=1}^{K} L_k} \qquad (8)$$

In the above function, $\delta(i, j)$ is the delta function, X_{kl} is the $l-$th speech feature vectors of the $k-$th speaker, and P_j is the observation probability of given feature sequence for m-th speaker. The parameter $\arg\max_j P_j$ is the index of the speaker having maximum probability and P_j is defined by equation (1). The optimization object function having optimization variable *f*, f_1 and f_2 expressed by equation (8) is nonlinear function and it is impossible to find the closed solution of the function. Thus the PSO approach [16] is applied for optimizing the object function.

4. Experiment and Investigation

4.1 Experimental database and specifications

For the experiments of this research work, the VidTimit audio-visual database was employed. The VidTimit database contains 43 speaker (19 female and 24 male), reciting short sentences selected from the NTIMIT corpus. The mean duration of each sentence is around 4 seconds, or approximately 106 video frames. The video of each person is stored as a sequence of JPEG images with a resolution of 512 × 384 pixels (columns × rows), with corresponding audio provided as a monophonic, 16 bit, 32 kHz WAV file.

4.2 Experiments and results

For our proposed method validation, we have used VidTimit database which contains 10 sentences for each speaker in audio and visual signal level. In this database, the audio signals are captured in the office environment noises. We have taken only 9 utterances of each speaker and divided into three data set (DS); Group I (1-3 utterances), Group II (5-7 utterances) and Group III (8-10 utterances). Group I (Gr. I) is used for speaker model generation while Group II (Gr. II) and Group III (Gr. III) are involved in PSO training and validation for f, f_1 and f_2 in testing stage for speaker identification. For creating mismatch between train and test signal, we have added babble noises (BN) to the test data which is obtained from NoiseX-92 database. For lip image lighting condition change, we have added an artificial illumination to the respective testing visual image as follows:

$$I(y,x) = w(y,x) + |\varphi| d + \delta \qquad (9)$$

where $y = 1,2,...,M_Y$, $x = 1,2,...,N_X$, and d is either y or x depending on the illumination direction, $\varphi = -\delta/\eta$ and δ is illumination in pixel. In our experiment we have added the artificial illumination from up-to-down (UD) and right-to-left direction. Examples of the region of interest (ROI) lip images with artificial illumination in mentioned direction are shown in Fig. 3.

(a)

(b)

(c)

Figure 3. Different images view with illumination changes (a) without illumination (b) with up-to-down illumination direction (c) with right-to-left illumination direction

For features dimension reduction, we have applied local PCA to both the audio and visual features vector during the noisy condition i.e., while we added the different noises and illumination to speech and lip images test data set respectively. However, we did not apply the local PCA for the database supplied speech data which are recorded in office environment having relatively high degree of signal distortion. Before features extraction we have transform the color image to gray image. Table 1 shows the overall specifications for our experiment.

TABLE I. GMM BASED EXPERTS SPECIFICATIONS

Modality	Specification
Audio	Features: MFCC + Delta Dimension reduction: local PCA Original Features Dimension : 38 No. of Principal Components : 10 Reduced Features Dimension : 10 No. of Mixtures in GMM: 3
Lip Image (ROI)	Input Image Resolution: 64×64 pixel Features: DCT Dimension reduction: local PCA No. of Principal Components: 10 No. of Mixtures in GMM: 3

TABLE II. Baseline Audio-visual Fused System Performances

Environment	Noise Types to Speech signal	Average Signal-to-noise ratio (SNR) (dB)	δ	Light Direction	m	Train DS Gr.	Test DS Group	IR (%)
1	Office	4.90	160	UD	$-\delta/80$	I	II+ III	91.86
2	Babble Noises	1.92	160	UD	$-\delta/80$	I	II+III	91.47
3	Office	4.92	160	RL	$-\delta/80$	I	II+ III	92.24
4	Babble Noises	1.90	160	RL	$-\delta/80$	I	II+III	91.47

TABLE III. Proposed Methods based Audio-visual Fused System Performances

Environment	Train DS	Test DS for PSO	Optimum f	Optimum f_1	Optimum f_2	Test DS by optimum f, f_1 & f_2	IR (%)	Avg. IR (%)
1	I	III	0.552	0.537	0.1497	II	96.89	96.89
	I	II	0.575	0.616	0.1431	III	96.89	
2	I	III	0.411	0.377	0.1728	II	93.02	93.80
	I	II	0.432	0.383	0.1462	III	94.59	
3	I	III	0.459	0.482	0.044	II	97.67	97.28
	I	II	0.394	0.424	0.051	III	96.89	
4	I	III	0.483	0.611	0.041	II	92. 24	93.4
	I	II	0.452	0.563	0.062	III	94.57	

The different experimental results are presented in Table II and Table III. Table II shows the experimental result of our adopted baseline system. We have presented our proposed method based experimental results in Table III as those of environment condition of Table II. It is mentioned in the previous section that for the experiment in environment 1 the audio features are not subjected by local PCA. It is seen from the Table III that the proposed optimized factor in correcting the missing reliability is validated through the values of different f and maintains the consistency in the fusion process and also the optimal f_1 and f_2 maintained the reliability statistics. The proposed methods based system fulfilled the desire of fusion strategy through an improvement of speaker identification rate of 5.02% and 2.33% in environment condition 1 and environment condition 2, respectively for up-to-down lighting condition. Similarly, for left-to-right direction, sufficient improvements have been depicted in Table III. Moreover, the proposed optimized factor in reliability correction protects the bimodal system not to transform into single modality system thus fulfils the desire of fusion strategy by taking both modalities information simultaneously.

5. Conclusion

In this paper, we have presented a robust bimodal person identification system considering the correction of reliability and its optimality in the integration process of audio and visual information. Introducing the adaptive threshold and optimization factors in the existing convection function improved the performance of the bimodal robust speaker identification system significantly. Moreover, the proposed adaptive threshold clearly protect the bimodal biometrics system not to undergo into single modality and optimization factors clearly balanced the trend of confidence level between the modalities in terms of modalities performances. The optimizing factors were optimized by PSO algorithm. The entire experiments were performed using the VidTimit database. With the bimodal speaker identification system, we confirm that the proposed adaptive threshold and the modified convection function could be a promising solution in multimodal biometrics technology. For further study, we will apply a range of different degree of noises and illumination condition using different models for text dependent speaker identification and specifically will focus on multimodal based speaker identification, audio-visual speaker verification and particularly the audio-visual speech recognition in real environment.

Acknowledgments

“This work is supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National Industry Promotion Agency)” (NIPA-2010-C1090-1011-0008).

References

[1] J. P. Campbell, “Speaker recognition: a tutorial,” Proc. IEEE, vol.85, no.9, pp. 1997, pp. 1437-1462.

[2] D. A. Reynolds, “An overview of automatic speaker recognition technology,” Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing, vol. 4, 2002, pp. 4072-4075.

[3] A.K. Jain, A. Ross, S. Prabhakar, “An introduction to biometric recognition,” IEEE Tran. Circuits Sys. Video Technol., vol. 14, no.1, 2004, pp. 4-20.

[4] D. A. Reynolds, R. C. Ross, “Robust text-independent speaker identification using Gaussian mixture speaker models,” IEEE Trans. Speech Audio Proc. vol. 3, no.1, 1995, pp. 72-82.

[5] B.S. Atal, “Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification,” J. Acoust. Soc. Am., vol. 55, no. 6, 1974, pp. 1304-1312.

[6] R. J. Mammone, X. Zhang and R. P. Ramachandran, “Robust speaker recognition: a feature-based approach,” IEEE Signal Processing Magazine, vol. 13, no. 5, 1996, pp. 58-71.

[7] B. Zhen, X. Wu, Z. Liu and C. Huisheng, “An Enhanced RASTA processing for speaker identification,” Proc. ICSLP, pp. 251-254, 2000.

[8] E. Mengusoglu, “Confidence measure based model adaptation for speaker verification,” Proc. 2nd IASTED Internat Conf. on Communications, Internet and Information Technology, 2000, pp. 408-411.

[9] U. V. Chaudhari, G.N. Ramaswamy, G. Potamiannos, and C. Neti, “Audio-visual speaker recognition using time-varying stream reliability,” Proceeding of IEEE Int. Conference on Acoustics, speech and signal proc. vol. 5, 2003, pp. 712-715.

[10] M. Heckmann, F. Berthommier, and K. Kristian, “Noise adaptive stream weighting in audio-visual speaker identification,” EURASIP J. Applied Signal Proc. 2002, pp. 1260-1273.

[11] N. A. Fox, Audio and Video Based Person Identification, Doctoral thesis, University College Dublin, 2005.

[12] T. Wark, and S. Sridharan, “Adaptive Fusion of speech and lip information for robust speaker identification,” Digital Signal Processing, vol. 11, 2001, pp.169-186.

[13] C. Sanderson, Biometric Person Recognition: Lip, Speech and Fusion. VDM-Verlag, 2008.

[14] M. Tariquzzaman, Jin Young Kim and Joon-Hee Hong, “Improvement of reliability based information integration in audio-visual person identification,” J. Korean Soc. of Phonetic Sci. Speech Technol. Vol. 62, 2007, pp. 149-161.

[15] S. Furui, “Cepstral Analysis technique for automatic speaker verification,” IEEE Trans. on Acoustics, Speech, and Signal Proc., vol. 29, no.2, 1981, pp. 254-272.

[16] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” Proc. Sixth Int. Symposium on Micro Machine and Human Science, 1995, pp. 39-43.

[17] N. Kambhatla, and T.K. Leen, “Dimension reduction by Local PCA,” Neural Computation, vol. 9, no.7, 1997, pp. 1493-1503.

[18] M. Tariquzzaman, Jin Young Kim and Seung You Na, “Robust Person Identification Using Optimal Reliability in Audio-Visual Information Fusion,” The J. of Acoustical Soc. Of Korea, vol. 28m No. 3E, 2009, pp. 109-117.

Md. Tariquzzaman received the B.Sc. and M.Sc. degree in Applied Physics, Electronics and Communication Engineering (Erstwhile Electronics & Applied Physics) from Islamic University, Kushtia-7003, Bangladesh, in 2001 and 2003, respectively. He joined as a lecturer in the Dept. of Information and Communication Engineering, Islamic University, Kushtia-7003, Bangladesh, in September 2004. He is currently pursuing PhD degree in the School of Electronics & Computer Engineering at Chonnam National University, Republic of Korea. His research interests are in the area of biometrics, computer vision and machine learning.

Kim Jin Young received the B.S., M.S. and Ph.D. degrees in Electronics Engineering from Seoul National University in 1986, 1988 and 1993. He is currently working as a Professor in the School of Electronics & Computer Engineering at Chonnam National University, Republic of Korea. His research interests are in the area of biometrics, audio-visual signal processing, and speech recognition.

Seung You Na received the B.S. degree from the Dept. of Electronics Engineering from Seoul National University in 1977. He received Ph.D. degree in ECE from University of Iowa, USA. He is currently working as a Professor in the School of Electronics & Computer Engineering at Chonnam National University, Republic of Korea. His research interests are in the area of intelligence control, signal processing,.

Mohd A. Rashid received the B.Sc. degree in Electrical and Electronic Engineering from Bangladesh Institute of Technology (BIT), Khulna in 1991. He also received Master’s and Ph.D. degrees in Electrical and Information Engineering from University of the Ryukyus, Japan in 2000 and 2003, respectively. He served Electrical and Electronic Engineering Department of BIT Khulna as a faculty from 1993 to 2005. He also worked as Visiting Researcher at Wilfried Laurier University and University of Waterloo Canada during 2005/2006. He is currently working at the University of the Ryukyus, Japan as JSPS Postdoctoral Research Fellow. His multidisciplinary research interests include wireless networks, computational electromagnetics, bioinformatics, modeling and simulation of real world problems.

Shiro TAMAKI received his B.S. degree from the University ofthe Ryukyus, Okinawa, Japan in 1979. He also received M.S. & Ph.D. degrees from the Graduate School of Engineering, University of Tokushima, Japan in 1983 and 1988, respectively. He is currently working as a professor at the Department of Information Engineering, University of the Ryukyus, Okinawa, Japan. His research interests include the Coding theory, Cryptographic theory, Environmental problems, Natural energy application and Optimization theory. He is a member of the Institute of Electronics Engineers of Japan, Japan Solar Energy Society, and Japan Association for Education Technology.