Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SPEAKER IDENTIFICATION AND LOCALIZATION USING FUSION OF FEATURES AND SCORE LEVEL FUSION

[1] RASHA H. ALI, [2] DR. MOHAMMED NAJM ABDULLAH AND [3] DR. BUTHAINAH F. ABED

[1] Lecturer in University of Baghdad- College of Education for Women- Computer department and Ph.D. Student in ICCI, Iraq

[2] ASST. PROF., Lecturer in Iraqi Commission for Computers and Informatics (ICCI) - Information Institute for Postgraduate Studies, Iraq

[3] ASST. PROF., University of Information Technology and Communications, Iraq

E-mail: [1]ha170642@gmail.com, [2] mustafamna@yahoo.com, [3]buthynna@yahoo.com

ABSTRACT

The localization and identification of speaker used in diverse application such as meeting, conferences, smart environments and robot-human interactions. So, the accuracy is perfectly significant of these systems which is increasing in the proposed system. In this paper the proposed system depends on identification and localization features. Three stages are presented: the preprocessing, the stage of extraction for the feature and the classification stage. In the stage of preprocessing, the energy and zero crossing techniques that are be using to split voice and silent of the speech signals. While in the stage of feature extraction, the fusion level features that are using for identification and for localization implemented with six features in both domains (the domain of time and frequency). For identification features: - the energy and the zero crossing were extracted in a time domain. The entropy feature was extracted after computation the wavelet transform. The spectral centroid, spread and spectral entropy were extracted after computation the Fourier transform. While for a localization features, the Capon beam forming (MVDR) was implemented. In a classification stage, the random forest was used and the score level fusion technique for random forest and the support vector machine. The ELSDSR dataset was used for training and testing, which contains 198 file sound. The accuracy of the system was 88.050% when using the random forest, and 95.226% when using score level fusion

Keywords: *Speaker Localization, Speaker Identification, Random Forest, Support Vector Machine, Feature Level Fusion, Score Level Fusion.*

1. INTRODUCTION

Communication is a multidimensional dynamic process that is necessary to express thoughts, emotions, and needs, allowing interaction between people and their environment. The hearing, motor coordination and speech production are included in the process of communication. If one of these aspects or more is impaired, the communication is disordered[1].

The analysis for Auditory Scene (ASA) is the basis in human communication. The capabilities of human auditory permit us to identify speakers and speech, localize and separate different speakers or sound sources[2]. The signal of speech includes significant paralinguistic information, like identity, gender, age, accent, the state of emotional for the speaker and language[3].

Speaker Recognition (SR) is the procedure for identifying a speaker according to the features of the vocal[4]. Speaker recognition (SR) contains two methodologies, firstly Speaker identification(SI) and secondly speaker verification(SV). SI is determining the identity of unknown speaker depend on his/her utterance properties. Speaker Identification(SI) is used in various application such as shopping using the telephone, voice dialing, mail of voice, and the services for access to database. Also, the speaker recognition (SR) is classified into two tasks: firstly, text-dependent and secondly the text-independent. In the text-dependent that is constrained, the phrase of speech for the speaker and the lexicon knowledge is incorporated into the modeling, while the text-independent isn't constrained the phrases of the speaker. The results of recognition in text-dependent are more precise compared to the systems of text-independent [5]. The system of human auditory has abilities to realize and localize a target source in complex multi-source scenarios. The speaker localization is an important technique for the ability of the robot to detect the directions of talkers for expression of the interest in the conversation by using speaker localization systems. So it remained the challenging function for algorithms that are used for localization[6].

1.1 The Contribution of The Work

The Ref [7-9] presents the identification and localization of the speakers using camera video for acquisition the information that used for tacking the person and detect the identity of the speaker, the researches that depend on the camera video used the audio and visual features that gave are good results but high cost. In Ref [10], he speaker localization and identification had been presented in two stages: localization and identification using MFCC. While the contribution of this work is determining the position of speakers and identification of a speaker in closed spaces in parallel, By using the techniques of speaker localization for determination the localization of a speaker and used these techniques for identification by concatenation the feature of identification and the feature of localization to construct the feature fusion using Radio-Frequency Identification (RFID) that used one or array of a microphone for data acquisition which depend only on the audio features. And because those systems using in many application and system such as the environments of smart, The interaction between the robot and human. So the improvements of the accuracy are very important and necessary, so the paper presents the novel method by using the score level fusion also known decision fusion gave more improvement in an accuracy.

2. THE DATASET

The dataset ELSDSR that was used for training and testing the system. The purpose of creating the speech database of English language for the recognition of speaker is to take out wealthy voice messages with respect to gauge inter and intra speaker variability. Subjects are enrolled in a Danish technical university environment. Most of them are non-native English speakers. It supplies a good rate for speaker recognition. ELSDSR contains voice messages from 22 speakers (12M/ 10F), and the age covered from 24 to 63. the speaker was different in distribution with respect of nationality and age, except for the gender[11].

3. THE PROPOSED SYSTEM

In general, the speaker identification contains the following stages: -

1. *Front-end processing (pre-processing):* - this step related to the “signal processing” part, which converts the continuous signals into discrete signals, removes the noise or background noise from signals using different techniques.
2. *Feature extraction:* - this step for extracting the acoustic features from signals of speakers to construct set of vectors of feature for use in training and testing phases.
3. *Decision-making*:- it’s the final decision step for an identity of a speaker by comparing unknown feature vectors to all models in a database[12]. Fig.1 shows the components of a proposed system, while the steps of the system shown in algorithm 1.

3.1. The Preprocessing Stage

In this stage contains the following steps:

a. Pre emphasize the signal, the idea of pre-emphasis is to spectrally flatten the speech signal and equalize the inherent spectral tilt in speech, it is used by the first order First Impulse Digital Filter (FIR) using the following equation

$$H_p(Z) = 1 - az^{-1} \quad (1)$$

Where a is a constant, which value is, $a = 0.97$ [13].

b. Framing the signal of speech into blocks with 25ms with overlap 10ms.

c. Windowing the speech signal by applying the hamming window with a view to keeping the signal continuity, so each frame multiply with hamming window using the following equation[14].

$$H(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \quad (2)$$

for n= 0,1,…,N-1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where n is a current sample, and N: number samples in each frame.

d. The detection of voice activity (VAD) was applied, VAD is the operation of setting the presence of human speech [4]. And, for splitting the signal of speech into voiced and unvoiced signal we used the energy of short time (STE)[15] and the rate of zero crossing (ZCR) for each frame[16] using the Eq. (3) and (4) respectively.

$$E_n = \sum_{n=-\infty}^{\infty} x^2(n) \quad (3)$$

Where, n is the current sample

$$ZCR = (1/T-1)\sum_{t=1}^{T-1} \mathbb{I}\{S_t S_{t-1} < 0\} \quad (4)$$

Where, s is a signal of length T and the indicator function $\mathbb{I}\{A\}$ is egalitarian to one When A is True and otherwise is egalitarian to 0. While, the indicating the speech segment by the STE and ZCR using the threshold for energy. Eq.(5) shows the threshold of energy[17].

$$Threshold = Emin + 0.1\,(Emax - Emin) \quad (5)$$

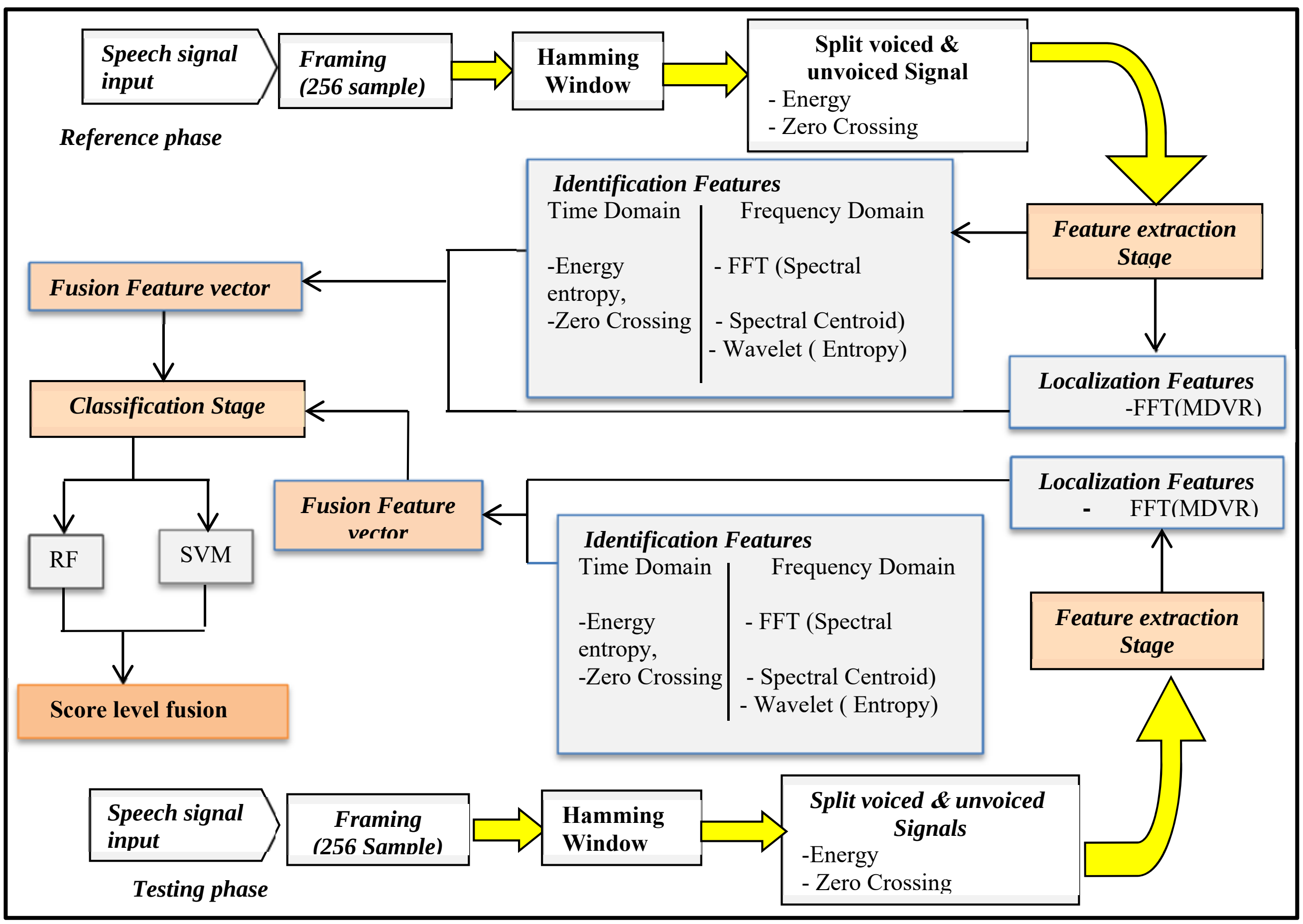


Figure 1: The proposed system

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Algorithm 1: The Speaker Identification and localization System

Input:- The speech signal.

Output:- The speaker identification.

Begin

Step1:- read the speech signal.

Step2:- pre- emphasize the signal using the Eq.(1).

Step3:- Framing the signal into blocks with 256 sample in each frame.

Step4:- Apply hamming window for each frame using Eq.(2).

Step5:- Voice Activity Detection(VAD) using the STE and ZCR by the Eq.(3) and(4) respectively and Eq.(5).

Step6:- Extract the feature using STE and ZCR in time domain using the Eq.(3) and (4).

Step7:- Compute Fourier transform for the signal using the Eq.(10).

Step8:- Extract feature using SE and SC using the Eq.(14) and (15) respectively.

Step9:- Extract feature for localization using the MDVR technique by the Eq.(20).

Step10:- Compute the Wavelet transform (Haar type) using the Eq.(8).

Step11:- Extract the Entropy using the Eq.(12).

Step12:- Classification the fusion features using RF by the Eq.(23).

Step13:- Classification the fusion features using SVM by the Eq.(24).

Step14:- Classification the fusion features using score level fusion for RF and SVM.

END

3.2. The Feature Extraction Stage

After the stage of preprocessing, the frames of speech signal were used for extraction the features for introducing the fusion features vector for identification and localization for both domains (time and the frequency domain) and features as show in the following step: - fig.2 displays the stage of the extraction for features.

a. The Short Time Energy (STE) : - compute the STE for the speech signal using the Eq.(3).

b. Zero Crossing Rate (ZCR) : - Compute the ZCR for the speech signal using the Eq. (4).

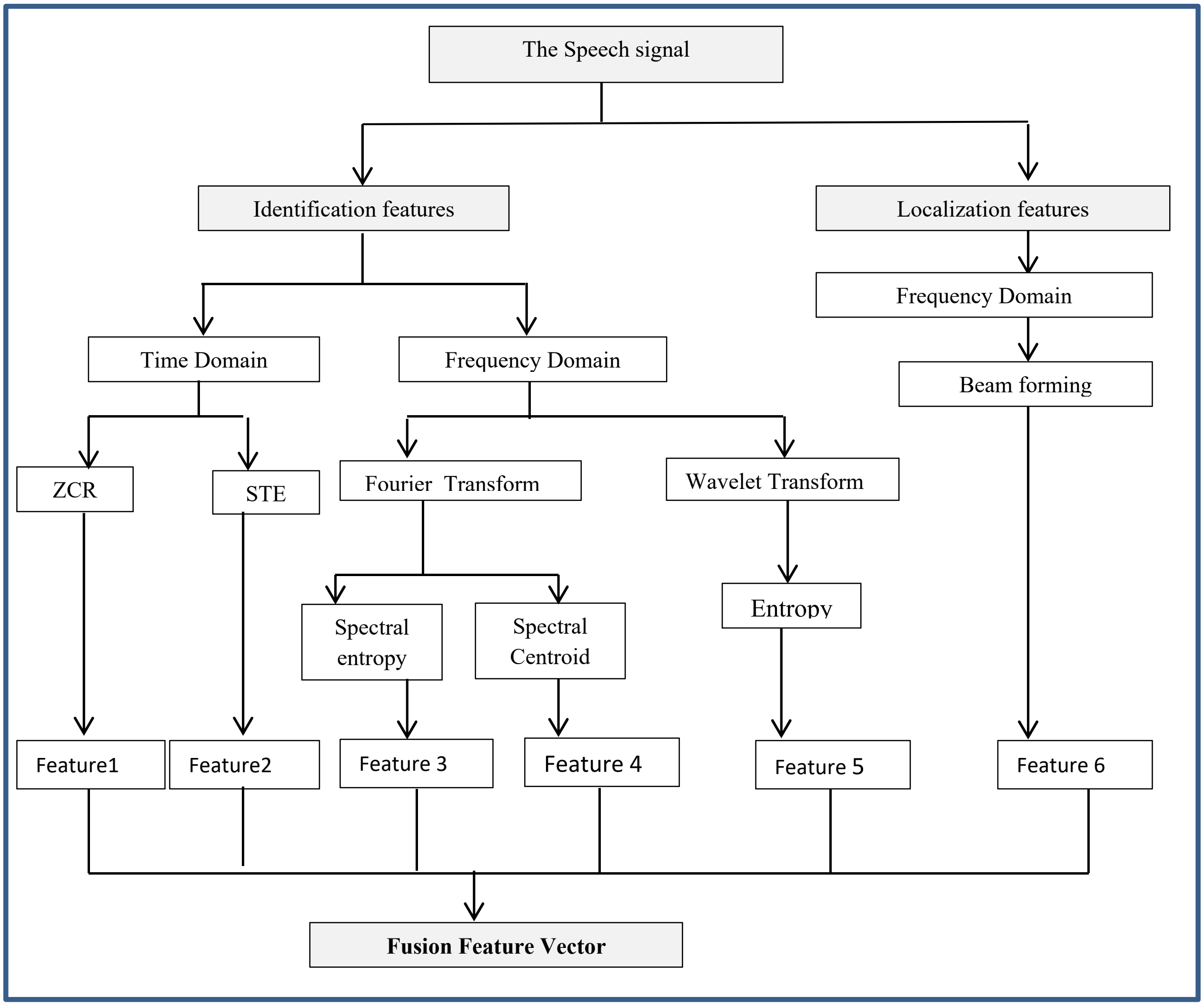


Figure 2: The feature extraction stage

c. *The Discrete Wavelet Transform (DWT):-*

Wavelets are efficient tools for converting the time into the domain of frequency and more effective tool for analysis of non-stationary signals[18]. The process of DWT is decomposing of the input signal into groups of function, this sets called wavelets. The DWT is accomplished by using mother wavelet function by the shifting and scaling. So the result of DWT process is wavelet coefficients. The signal has been reconstructing as the linear collection of wavelets and coefficients of weighting wavelet. After every decomposition, the coefficients that are resulted from the divided of the signal are approximation coefficients A and detailed coefficients D[19].

The transform of continuous wavelet (CWT) for the signal $f(t)$ in time t is defined in the following eq.(6)[20]

$$wt(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \Psi^{*}\left(\frac{t-b}{a}\right) dt \qquad (6)$$

Where $\psi_{a,b}(t) = \psi((t-b)/a)\big/\sqrt{a}$ with $a,b \in R$, and $\psi^{*}(t)$ is complex conjugate of the mother function $\psi(t)$. a is scale parameter and b is translation parameter. For the discrete parameter, $a = 2^{-j}$ and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$b = 2^{-j}k$, with $j,k \in Z$ (set of integers), and the base wavelet family is then expressed as:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \qquad (7)$$

As a result, the DWT for the signal $f(t)$ is obtained as:

$$wt(j,k) = \langle f(t), \psi_{j,k}(t) \rangle = 2^{j/2}\int_{-\infty}^{\infty} f(t)\psi^*(2^j t - k)dt \qquad (8)$$

where the symbol $\langle . \rangle$ denotes the operation of inner product [21].

d. Fourier Transform(FT): -

The transform of Fourier of a signal is called the signal representation in the domain of the frequency. The transform of Discrete Fourier (DFT) are quite useful because they detect periodicities for input data. The limitation condition in DFTs is that the series of discrete time must be periodic, and frequency of sampling should be great than twice of the band-limited frequency.

The discrete function case is $f(t) \rightarrow f(t_k)$ by letting $f_k = f(t_k)$, where $t_k = K\Delta$ with $K = 0,1,...,N-1$. Writing this gives the transform of discrete Fourier [22]:

$$F_n = F_k\left[\{f_k\}_{n=0}^{N-1}\right](n) \qquad (9)$$

As $$F_n = \sum_{k=0}^{N-1} f_k e^{-\frac{2\pi ink}{N}} \qquad (10)$$

The inverse transform

$$f_n = \frac{1}{N}\sum_{k=0}^{N-1} F_n e^{\frac{2\pi ink}{N}} \qquad (11)$$

In a DET, the signal is transformed from a domain of time (f_k) into a domain of frequency (F_n) [23]. For calculating the Eq.(9) or (10). The FFT is used for decreasing number of computations necessary for N points from $2N^2$ to $2N\log N$ where log is the base 2 logarithm [24].

e. The Shannon Entropy (SE): -

Entropy is the measure of quantitative of how uncertain the result of a random experience [25]. The Shannon's entropy is the suitable metric to gauge the organization and information sources of the signal. For signal x, the Shannon entropy is defined in Eq. (12)[26]:-

$$H = -\sum_{i=1}^{N} p_i \log p(i) \qquad (12)$$

Where p_i is the likelihood that the signal belongs to a considered period, with understanding that $p_i \log p(i) = 0$ if $p_i = 0$, and N is a number samples. The entropy H is a gauge of the information needed to determine a system in a specific case, the sense that H is the measure of our obscurity about the system.

f. Spectral Entropy (SE) : -

The spectral entropy the has been computed in the domain of frequency, after computing the FFT, then compute the Density of Power Spectral (PSD) using Eq.(13)[27, 28]. Then finally, compute the spectral entropy using the standard entropy equation as mention in Eq. (12).

$$p(w_i) = \frac{1}{N}|X(w_i)|^2 \qquad (13)$$

g. Spectral Centroid and Spread (SC and SS): -

The spectral centroid of the signal is the curve which the value at any given time is the centroid of the identical fixed-time cross portion of the signal's spectrogram[29]. The Eq. (14) is the spectral centroid equation.

$$C_t = \frac{\sum_{k=K_s}^{K_e} M_t[k] \times k}{\sum_{k=K_s}^{K_e} M_t[k]} \qquad (14)$$

Where C_t represents the spectral centroid of number for frame t, while $M_t[k]$ is the power spectrum with frequency bin number k, K_s is a lower limitation of frequency bandwidth bin number and K_e is a higher one[30]. The spectral spread (SS) is the measure of the average spread of the spectrum in relation to its centroid using the Eq. (15).

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$SS = \sqrt{\frac{\sum_{k=0}^{N/2} (f_k - SC)^2 |X(k)|^2}{\sum_{k=0}^{N/2} |X(k)|^2}} \quad (15)$$

h. Capon beamforming (Minimum Variance distortion less response (MVDR))

The minimum variance distortionless response (MVDR) or calledCapon beamformer is one of the direction of arrival (DOA) technique, which vastly used in sensor array signals processing applications. It depends on the estimation of the power of a signals that have multiple narrowband sources[31]. The MVDR is one of the techniques of adaptive beamforming that it is used to reduce a variance of the noise. The broadband MVDR beamformer divides wideband signals into narrowband components and treats these components independently[32]. The MVDR algorithm relies on the steering vectors. MDVR calculate the weight vector to locate the desired signal from the interference, maximize the sensitivity in one direction only and minimizes the power of output. The result is given at any time by the linear series of the data at source signals, with being the weight vector in Eq. (16)

$$y(n) = w^H(n)^*(n) \quad (16)$$

The Weight vector $W(n)$ is in the Eq.(17)

$$w(n) = \sum_{N=0}^{M-1} w_n, \quad x(n) = \sum_{n=0}^{M-1} X_n \quad (17)$$

the matrix inverse operation and uses the immediate gradient vector $\nabla J(n)$ for weight vector upgrading, the weight vector at time $n+1$ can be written as Eq.(18)

$$w(n+1) = W(n) + \frac{1}{2}\mu[\nabla J(n)] \quad (18)$$

Where μ is the step size parameter. While the Eq. (19) shows the weight vector

$$W(n+1) = W(n) + \mu[p(n) - R(n)W(n)] \quad (19)$$

$$= W(n) + \mu X(n)\left\lfloor d^*(n) - X(n)W(n) \right\rfloor$$

$$= W(n) + \mu Xe^*(n)$$

While R is covariance matrix, p is cross-correlation vector, $\nabla J(n)$ is gradient vector While the desired signal can be defined by the following equations:

$$y(n) = w^H(n)x(n) \quad (20)$$

$$e(n) = d(n).y(n)W(n+1)$$

$$= W(n) + \mu X(n)e^*(n)$$

The MDVR minimize the noise and high the signal noise ratio (SNR) using the Eq.(21)[33]

$$SNR = 10\log_{10}(\frac{S}{N}) \quad (21)$$

While S is the desired signal, and N is noise signal.

3.3 The Classification stage

In the classification stage, the feature vectors have been classified using score level fusion by using two classification techniques random forest (RF) and support vector machine (SVM) as following:

a. Random Forest (RF):

It is the ensemble of classification and regression tress (CART) [34]. RF is the technique used for regression and classification. It is trained on the same size of the datasets, a training set, it called bootstraps, created by resampling randomly the training group itself. Once a tree is built, the set of bootstraps, which do not include any special record from the genuine dataset out-of-bag (OOB) samples, it used as the test set [35]. In the algorithm of random forest, the random vector θ_k is produced, independent from the previous random vectors and distributed to all trees, and each tree is mature using random vector θ_k and training set, that results in the aggregate of classifiers for tree-structured $\{h(x,\theta_k), k = 1....\}$ at input vector x. In RF, the generalization error that estimate by the classification error as defined in Eq. (22)

$$PE^* = P_{X,Y}(mg(X,Y) < 0) \quad (22)$$

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where the random vectors are (X and Y) and mg is the margin function that measures that range to which the mean number of votes in random vectors for the right output override the mean vote for any other output. Margin function is defined as the Eq. (23)

$$mg(X,Y)=av_k I(h_k(x)=Y)-\max_{j\neq Y} av_k I(h_k(X)=j) \quad (23)$$

Where $I(.)$ is the indicator function[36].

b. Support Vector Machine (SVM) :

SVM is a technique used for classification, regression, and preference (or ranking) learning. The SVM can be formed in different forms (single class and multiclass). In single-class classification, The SVM as binary classifier where the result of the learned function is either negative or positive. In multiclass classification, the SVM is merging manifold binary classifiers using pairwise coupling method[37]. SVMs is one of a kernel methods category, that used in feature space of high dimensional for computing a dot product [38].

SVM classification:

$$\min_{f,\xi_i} \|f\|_K^2 + C\sum_{i=1}^{l}\xi_i \quad y_i f(\mathbf{x}_i) \geq 1 - \xi_i,$$

for all i $\xi_i \geq 0$ (24)

Dual formulation:

$$\min_{\alpha_i} \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j K(\mathbf{x}_i,\mathbf{x}_j)$$

$0 \leq \alpha_i \leq C$, for all i; (25)

Variables ξ_i are slack variables, they measure an error made in point ($\mathbf{x}_i$,y_i). Training SVM becomes quite challenging when the number of points for training is large[39].

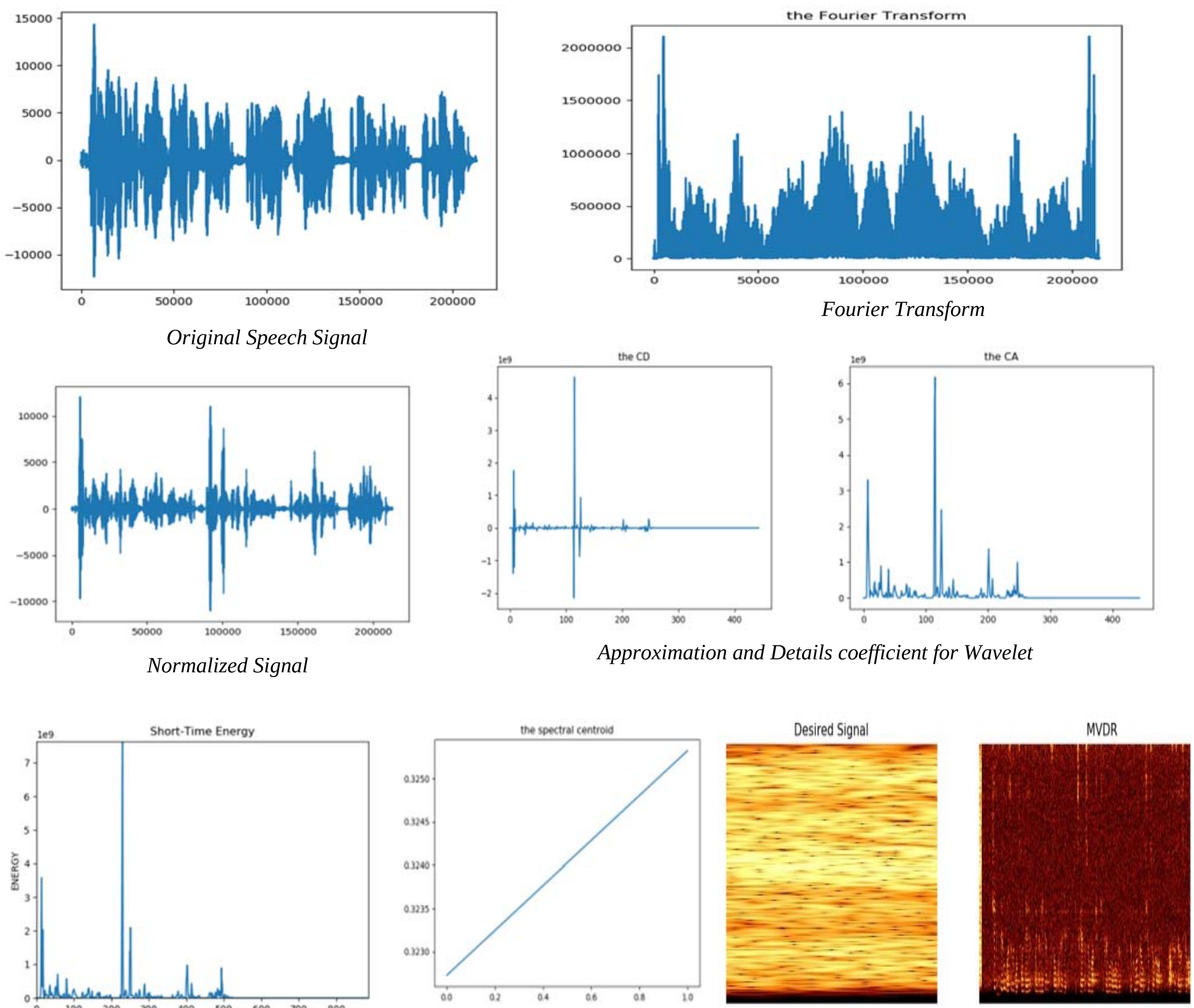
4. RESULT AND IMPLEMENTATION

The proposed system has been implemented on ELSDSR dataset with sampling rate 16000, 22 speakers (12M/10F). The dataset was divided into 80% for training and 20% for testing. The accuracy of the system was 88.05% when used random forest algorithm, while the accuracy was 95.22% when used score fusion level in classification for random forest and multiclass support vector machine. In fig.3. and Table 1. shows the features of a speech signal in the proposed system.

Table 1: The features of one for speech signal

The Name	*Total Sample*	*No. of frame*	*STE*	*ZCR*	*Entropy*	*SC*	*SS*	*SE*	*SNR*
Femal-Sp1	74400	310	198816.304	0.1221	3.0330	0.3227	0.3253	3.1025	20.047

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



Original Speech Signal *Fourier Transform*

Normalized Signal *Approximation and Details coefficient for Wavelet*

STE *Spectral Centroid and Spread* *Localization Using MDVR*

Figure 3: The results of system for one of speech signal

5. CONCLUSION

The importance of the speaker identification and localization systems accuracy is significant for identify and localize the person perfectly. This paper presented a system for increasing the accuracy of such systems by using feature vector fusion, score level fusion and one of a localization technique because of its capability for detecting the source of the speaker signal. A feature extraction phase is an important stage for identification process, therefore using six features (identification and localization features) with different domains to extract the good fusion feature vector. The results showed which there is an improvement in accuracy when using score level fusion and using RF and SVM in classification phase rather than using the one classifier. So the accuracy of these systems is significant for localizing and identify the person perfectly.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES:

[1] K. L. Kadi, S. A. Selouani, B. Boudraa, and M. Boudraa, "Fully automated speaker identification and intelligibility assessment in dysarthria disease using auditory knowledge," *Biocybernetics and Biomedical Engineering,* vol. 36, 2016, pp. 233-247.

[2] M. Li, J. Kim, A. Lammert, P. K. Ghosh, V. Ramanarayanan, and S. Narayanan, "Speaker verification based on the fusion of speech acoustics and inverted articulatory signals," *Computer speech & language,* vol. 36,2016, pp. 196-211.

[3] S. Safavi, M. Russell, and P. Jančovič, "Automatic Speaker, Age-group and Gender Identification from Children's Speech," *Computer Speech & Language,* 2018.

[4] S. S. Tirumala, S. R. Shahamiri, A. S. Garhwal, and R. Wang, "Speaker identification features extraction methods: A systematic review," *Expert Systems With Applications,* vol. 90, 2017, pp. 250-271.

[5] Y. Liu, Y. Qian, N. Chen, T. Fu, Y. Zhang, and K. Yu, "Deep feature for text-dependent speaker verification," *Speech Communication,* vol. 73, 2015, pp. 1-13.

[6] D. A. Blauth, V. P. Minotto, C. R. Jung, B. Lee, and T. Kalker, "Voice activity detection and speaker localization using audiovisual cues," *Pattern recognition letters,* vol. 33, 2012, pp. 373-380.

[7] Y. Hu, "Multimodal speaker localization and identification for video processing," *HKU Theses Online (HKUTO),* 2014.

[8] A. A. Salah, R. Morros, J. Luque, C. Segura, J. Hernando, O. Ambekar, *et al.*, "Multimodal identification and localization of users in a smart environment," *Journal on Multimodal User Interfaces,* vol. 2, 2008, pp. 75-91.

[9] K. Youssef, K. Itoyama, and K. Yoshii, "Simultaneous Identification and Localization of Still and Mobile Speakers Based on Binaural Robot Audition," *JRM,* vol. 29, 2017, pp. 59-71.

[10] H. Tómasson, "Speaker localization and identification," Citeseer, 2012.

[11] L.Feng,"Speaker Recognition," in *Informatics and Mathematical Modelling, echnical University of Denmark, DTU*, ed, 2004.

[12] C. Turner and A. Joseph, "A wavelet packet and mel-frequency cepstral coefficients-based feature extraction method for speaker identification," *Procedia Computer Science,* vol. 61, 2015, pp. 416-421.

[13] M. Kumar and A. G. Singh, "Performance analysis of LPC and MFCC techniques in automatic speech recognition," 2015.

[14] S. Yang, *Statistical approaches for signal processing with application to automatic singer identification*: Rochester Institute of Technology, 2016.

[15] R. C. Guido, "A tutorial on signal energy and its applications," *Neurocomputing,* vol. 179, 2016, pp. 264-282.

[16] K. Sundararajoo, "Improvement of audio feature extraction techniques in traditional Indian string musical instrument," Universiti Tun Hussein Onn Malaysia, 2015.

[17] E. Chandra and K. M. M. Kalaivani, "A study on speaker recognition system and pattern classification techniques," 2014.

[18] P. Fletcher and S. J. Sangwine, "The development of the quaternion wavelet transform," *Signal Processing,* vol. 136, 2017, pp. 2-15.

[19] E. Alickovic, J. Kevric, and A. Subasi, "Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction," *Biomedical Signal Processing and Control,* vol. 39, 2018, pp. 94-102.

[20] M. Tayyab, J. Zhou, R. Adnan, and X. Zeng, "Application of Artificial Intelligence Method Coupled with Discrete Wavelet Transform Method," *Procedia Computer Science,* vol. 107, 2017, pp. 212-217.

[21] X. Zhang, N. Feng, Y. Wang, and Y. Shen, "Acoustic emission detection of rail defect based on wavelet transform and Shannon entropy," *Journal of Sound and Vibration,* vol. 339, 2015, pp. 419-432.

[22] C. Cerovecki and S. Hörmann, "On the CLT for discrete Fourier transforms of functional time series," *Journal of Multivariate Analysis,* vol. 154, 2017, pp. 282-295.

[23] R. Kluger, B. L. Smith, H. Park, and D. J. Dailey, "Identification of safety-critical events using kinematic vehicle data and the discrete fourier transform," *Accident Analysis & Prevention,* vol. 96, 2016, pp. 162-168.

[24] B. Mukherjee, V. Srivardhan, and P. Roy, "Identification of formation interfaces by using wavelet and Fourier transforms," *Journal of Applied Geophysics,* vol. 128, 2016, pp. 140-149.

[25] O. A. Olvera-Guerrero, A. Prieto-Guerrero, and G. Espinosa-Paredes, "Non-linear boiling water reactor stability with Shannon

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Entropy," *Annals of Nuclear Energy,* vol. 108, 2017, pp. 1-9.

[26] R. Gu, "Multiscale Shannon entropy and its application in the stock market," *Physica A: Statistical Mechanics and its Applications,* vol. 484, 2017, pp. 215-224.

[27] T. P. Shri and N. Sriraam, "Spectral entropy feature subset selection using SEPCOR to detect alcoholic impact on gamma sub band visual event related potentials of multichannel electroencephalograms (EEG)," *Applied Soft Computing,* vol. 46, 2016, pp. 441-451.

[28] A. Abdelsamie, G. Janiga, and D. Thévenin, "Spectral entropy as a flow state indicator," *International Journal of Heat and Fluid Flow,* vol. 68, 2017, pp. 102-113.

[29] N. Kamarudin, S. Al-Haddad, S. J. Hashim, M. A. Nematollahi, and A. R. B. Hassan, "Feature extraction using spectral centroid and mel frequency cepstral coefficient for Quranic accent automatic identification," in *Research and Development (SCOReD), 2014 IEEE Student Conference on*, 2014, pp. 1-6.

[30] T. Furoh, T. Fukumori, M. Nakayama, and T. Nishiura, "A study of degraded-speech identification based on spectral centroid," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 2014, pp. 2946-2951.

[31] K. Elkhalil, A. Kammoun, T. Y. Al-Naffouri, and M.-S. Alouini, "Fluctuations of the SNR at the output of the MVDR with regularized Tyler estimators," *Signal Processing,* vol. 135, 2017, pp. 1-8.

[32] Q. Huang, R. Hu, and Y. Fang, "Real-valued MVDR beamforming using spherical arrays with frequency invariant characteristic," *Digital Signal Processing,* vol. 48, 2016, pp. 239-245.

[33] T. S. Kiong, S. B. Salem, J. K. S. Paw, K. P. Sankar, and S. Darzi, "Minimum variance distortionless response beamformer with enhanced nulling level control via dynamic mutated artificial immune system," *The Scientific World Journal,* vol. 2014, 2014.

[34] Z. Bei, Z. Yu, N. Luo, C. Jiang, C. Xu, and S. Feng, "Configuring in-memory cluster computing using random forest," *Future Generation Computer Systems,* vol. 79, 2018, pp. 1-15.

[35] A. Sarica, A. Cerasa, and A. Quattrone, "Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review," *Frontiers in Aging Neuroscience,* vol. 9, 2017, p. 329.

[36] Z. Masetic and A. Subasi, "Congestive heart failure detection using random forest classifier," *Computer methods and programs in biomedicine,* vol. 130, 2016, pp. 54-64.

[37] M. Li, X. Liu, and X. Liu, "Infrasound signal classification based on spectral entropy and support vector machine," *Applied Acoustics,* vol. 113, 2016, pp. 116-120.

[38] T. T. Ramanathan and D. Sharma, "Multiple Classification Using SVM Based Multi Knowledge Based System," *Procedia Computer Science,* vol. 115, 2017, pp. 307-311.

[39] S. Deb and S. Dandapat, "Classification of speech under stress using harmonic peak to energy ratio," *Computers & Electrical Engineering,* vol. 55, 2016, pp. 12-23.