WALIA journal 30(S2): 56-62, 2014
Available online at www.Waliaj.com
ISSN 1026-3861

Data reduction on MFCC features based on kernel PCA for speaker verification system

Mohd Azha Mohd Saleh *, Noor Salwani Ibrahim, Dzati Athiar Ramli

Intelligent Biometric Group (IBG), School of Electrical & Electronic Engineering, USM Engineering Campus, 14300, Nibong Tebal, Penang Malaysia

Abstract: Introduction: Mel Frequency Cepstrum Coefficients (MFCC) is one of the most widely used feature extraction techniques for speech recognition and produce MFCC features as input for the classification phase. In this study the reduction of feature dimension on MFCC features is studied due to large data size affects computational time which leads to slower verification speed. So, implementation of data reduction techniques so as to retain the most important feature parameters is evaluated in this study. In this study, an investigation of data reduction based on principal component analysis is proposed. Two approaches of Kernel Principal Component Analysis (KPCA) techniques i.e. Gaussian and Polynomial KPCA and PCA are evaluated and compared. The features based on MFCC and the reduced dimensions based on KPCA and PCA are then classified using two types of Support Vector Machine (SVM) classifiers i.e. linear and polynomial SVM. A set of clean data samples with three different dimensions of principle components i.e. 80, 117 and 180 are used for system evaluation. For performance evaluation, Equal Error Rates (EER) and verification time (VT) are employed in this study. The best system performance is observed for MFCC-KPCA Gaussian feature extraction technique with 117 features dimensions using linear SVM as classifier. This study proves that the use of data reduction technique can speed up verification time tremendously and improve system performances as well.

Key words: *Data reduction; MFCC; Kernel PCA; Speaker verification*

1. Introduction

Speech recognition systems have been developed successfully by utilizing many types of feature extraction methods. One of the most popular methods is Mel Frequency Cepstrum Coefficients (MFCCs) which provide a compact parametric representation of a cosine transform of the real logarithm of the short-term energy spectrum expressed on a Mel frequency scale. In 1980, Davis and Mermelstein (1980) compared parametric representations for monosyllabic word recognition in continuously spoken sentences. This study utilized parametric representation based on Mel Frequency Cepstrum, Linear Frequency Cepstrum, Linear Prediction Cepstrum and Linear Prediction Spectrum. Based on the experimental results, this study concluded that MFCC possess significant advantage over the other methods. The superior performance of the MFCC may be attributed by the fact that they are better in representing the perceptually relevant aspects of the short-term speech spectrum.

Lee, Fang, Hung and Lee (2001) have presented a new feature extraction approach that designs the shapes of the filters in the filter bank. The study applied PCA approach on the FFT spectrum of the training data. As a result the conventional MFCC features have been improved by the PCA-optimized filter bank. The proposed features are robust to additive noise for speech recognition while providing the same result for clean speech. It is claimed due to the PCA-optimized filter bank has maximized both the SNR variance ratio and the variation of the features.

A comparative evaluation on various MFCC implementations has been implemented by Ganchev et al. (2005). The implementation differs from other researches mainly in the number of filters, the shape of the filters, the way the filters are spaced, the bandwidth of the filters, and the manner in which the spectrum is warped. In addition, the frequency range of interest, the selection of actual subset and the number of MFCC coefficients employed in the classification are also evaluated. As a result, this study reported that the speaker verification performance does not vary vastly when different approximations of the non-linear pitch perception of human are used. However, some observations suggested that regardless of the specific filter bank design, a larger number of filters favour the speaker detection performance. Beside the number of filters in the filter bank, the overlapping among the neighbouring filters also proved as a sensitive parameter.

Chen and Luo (2009) in their paper have proposed a study on the use of MFCC and SVM for text-dependent speaker verification. The MFCCs used in this paper are extracted from the voiced password spoken by the user. These parameters are

* Corresponding Author.

then normalized and then used as the speaker features for training a claimed speaker model via SVM. By using speech signals selected from the Aurora-2 database, experimental results shown the performance of the proposed speaker verification algorithm yields an average accuracy rate of 95.1% with 22-order MFCCs.

Although speech recognition systems have been developed successfully with great performances and features, this success depends much on the extracted speech features, which has an important role in the whole recognition system (Amaro et al; 2004). If the speech features are not well extracted or come with an extreme data size, it will cost much computational time to the speaker verification system, which then will affect its performance and speed.

Further research on MFCC applying data reduction in speaker recognition system has been done by Hasan, Jamil and Rahman (2004). This study presented a security system based on speaker identification by utilizing MFCC as feature extraction method while Vector Quantization (VQ) technique as data reduction method. The study revealed that, when the number of centroids increased, the identification rate of the system also increases. They also found that combination of Mel frequency and Hamming window leads to better performance. This study also observed that the linear scale can improve system accuracy if comparatively higher number of centroids is used. However, the recognition rate using a linear scale would be much lower if the number of speakers increased.

The use of MFCC and Vector Quantization in speaker recognition has also been carried out by Mishra and Agrawal (2012). This study implemented an enhanced MFCC with silence removal. The silence present before and after the voiced part is removed to improve the performance of classifier. Based on their findings, this research suggested an effective normalization algorithm can be adopted on extracted parametric representations in order to improve the identification rate. Apart from that, a combination of features i.e. MFCC, LPC, LPCC, Formant etc. may be used so as to obtain a robust parametric representation for speaker identification.

The combination of MFCC and PCA has also been presented by Ittichauchareon, Suksri and Yingthawornsuk (2012). This paper described an approach of speech recognition using the MFCC features extracted from speech signal of spoken words. PCA is employed as the supplement in feature dimensional reduction state, prior to training and testing speech samples via Maximum Likelihood Classifier (ML) and SVM. It is found that the combination of MFCC-PCA-SVM with more MFCC samples have shown the improvement in recognition rates significantly compared to MFCC-PCA-ML.

Motivated by all of these researches, this study comes out with the proposed system of MFCC-KPCA-SVM model for speaker verification system where KPCA is used as data reduction technique on MFCC features. PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences (Rodriguez, de Paz et al; 2008). Since patterns are hard to find in high dimensional data, PCA is a powerful tool for analyzing data.

According to Leitner, Pernkopf and Kubin (2011), linear PCA refers to orthogonal transformation of the space containing the data samples. The transformed space is spanned by the eigenvectors that are found by eigenvalue decomposition of the covariance matrix estimated from the data samples. The coordinates of the data samples after transformation are referred as principal components. Normally, few principal components capture most of the characteristics of the data. The directions of these components are given by the eigenvectors corresponding to large eigenvalues, as a large eigenvalue means that its eigenvectors covers relevant information of the data.

Kernel PCA (KPCA) is one of the kernel algorithms that have been known from mid-nineties. It performs the PCA in the feature space, so it looks for directions of largest variance that yields nonlinear directions in the input space. KPCA was introduced after PCA to merit the performance of PCA (Huang et al; 2009). KPCA is a non-linear extension of PCA which data is first mapped and PCA is applied to the mapped data. KPCA make it possible for us to represent the speech features in a higher dimensional space which can possibly generate more distinguishable speech features (Amaro et al; 2004). It can extract up to *n* (number of samples) nonlinear principal components without expensive computations. It also can give a good re-encoding of the data when it lies along a non-linear manifold.

KPCA involves calculation of the eigenvalues decomposition or singular value decomposition of centered kernel data and is in search for orthogonal functions that optimize the kernel data scatter. In the linear case, it is well known that the classical PCA is not robust against data contamination and a small portion of outliers can give disturbance to the resulting principal components (Huang et al; 2009). PCA is a powerful technique for extracting structure from possibly high-dimensional data sets. But it is not effective for data with nonlinear structure. In KPCA, the input data with nonlinear structure is transformed into a higher dimensional feature space with linear structure, and then linear PCA is performed in the high-dimensional space.

So, in next section, discussion on feature extraction method using MFCC is presented. Subsequently, PCA and KPCA as data reduction techniques are described. Finally, explanation on SVM classifier as classification method is then given.

2. Material and methods

2.1. Research framework

The overall research framework of this study is summarized as in Fig. 1 below.

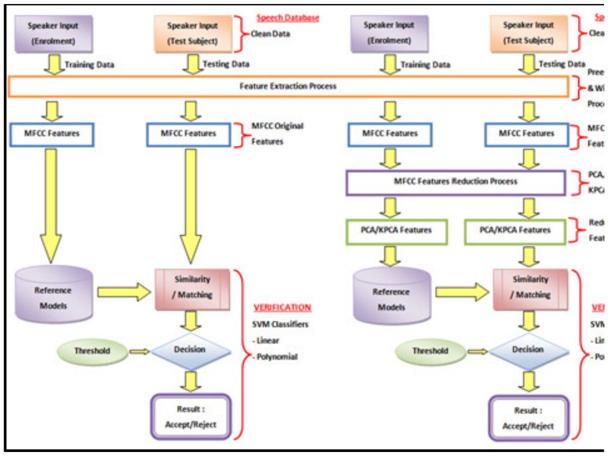

Fig. 1: Research framework

2.2. Database

The digitized audio signals from the Audio Visual Digit Database (Sanderson and Paliwal; 2001); which is monophonic, 16 bit, 32 kHz in WAV format have been used for performance evaluation in this research. This database consists of video and corresponding audio recordings of 37 speakers (21 males and 16 females). The recordings are done in three sessions. In each session, each speaker performed 20 repetitions of digit zero to nine hence 60 audio data for each speaker from all sessions. In total, 2220 audio data from entire speakers have been used for this research.

2.3. Feature Extraction

The entire process of extracting MFCC features is illustrated in Fig. 2. In this research, we utilize a feature set consists of 12 mel cepstrum coefficients, one log energy coefficient, 13 delta coefficients and 13 delta^2 coefficients per frame which in total 39 coefficients. The entire frames are then resized with data interpolation technique to 64×64 matrix. This feature matrix is then reshaped to 1×4096 as feature vector to represent each voice sample. A matrix of 740×4096 dimensions based on 20 voice samples and 37 speakers is then constructed.

For the purposes of comparison and evaluation with the proposed method of PCA and KPCA, we resize the feature matrix of 64×64 dimensions to 10×10, 12×12 and 14×14 using data interpolation technique. This entire new matrix is then reshaped to 1×100, 1×144 and 1×196 respectively which will represent each voice samples. As a results, for training data with 20 voice samples and 37 speakers, four sets of feature dimensions i.e. 740×100, 740×144, 740×196 and 740×4096 are used.

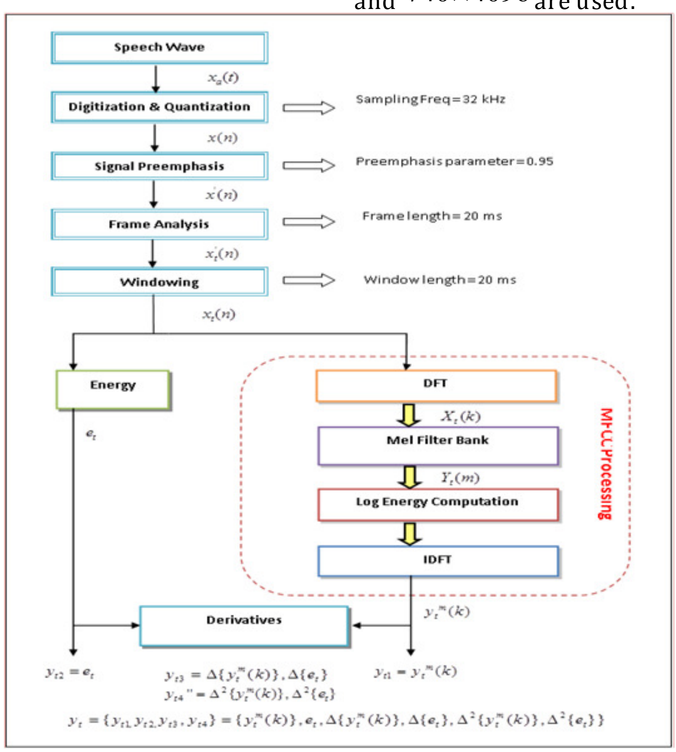

Fig. 2: Feature extraction process

3. Principal Component Analysis (PCA) Processing

For PCA technique, a set of eigenvoices (eigenvectors) from the training data is constructed. This eigenvoices are then used to generate the projection of new training and testing data. Our training data consist of MFCC features of matrix in $N^2\times M$ dimension where N^2 is the feature sizes and *M* is sample size hence *x* = 4096×740.

The eigenvoices for training data are computed as the following steps.

Step 1: Computation of mean and subtraction of mean from each data point of all samples.

Define the data average vector:

$$\mu = \frac{1}{M}\sum_{i=1}^{M} x_{ij} \quad (1)$$

where x_{ij} represent each data point in matrix x for i=1,2,3,...,M and $j = 1, 2, 3, \ldots, N^2$

Step 2: Compute the covariance matrix

$$XX^T = \frac{1}{M-1}\sum_{i=1}^{M}(x_{ij} - \mu)(x_{ij} - \mu)^T \quad (2)$$

Determine the Eigenvoices and Eigenvalues

In order to get the eigenvectors and eigenvalues, we need to compute a covariance matrix which characterizes the distribution of all samples.

$$C = XX^T \quad (3)$$

By using equation (3), a very large matrix with the dimension of $N^2 \times N^2$ is produced i.e. 4096×4096. The large size matrix will also add a large computational burden for the analysis. This can be avoided by using an optimum method introduced by Turk and Pentland (1991) as below:

$$C = X^T X \quad (4)$$

where a smaller matrix of $M \times M$ dimensions (740×740 in this research) is used.

Then, the Eigenvector φ_i and Eigenvalue ψ_i can be calculated as below:

$$C\varphi_i = \psi_i \varphi_i \quad \text{for } i=1,2,3,\ldots,M \quad (5)$$

$$\psi_i = \frac{1}{M}\sum_{j=1}^{M}\left(\varphi_i^T \varphi_j\right) \quad (6)$$

while the eigenvalue in equation (6) is subjected to the following conditions:

$$\varphi_j^T \varphi_i = \begin{cases} 1 & i = j \\ 0 & elsewhere \end{cases} \quad (7)$$

Step 3 : Project the original data into eigenspace.

Then the Eigenvoices is created by multiplying matrix A with each column of the Eigenvector above using the following equation,

$$v_k = A\varphi_k \quad (8)$$

where φ_k is the Eigenvector column with k[th] elements. As a result we get the Eigenvoices as follows,

$$V = \left[v_1, v_2, v_{3,\ldots}, v_Q \right] \quad (9)$$

The size of Eigenvoice matrix, V is $N^2 \times Q$, as a result of multiplying A with dimension of $N^2 \times M$ and φ_k with dimension of $M \times Q$.

As the Eigenvoice matrix, V has been created; the projection of training data into the eigenspace can be done. The projection is implemented using the equation below,

$$Y_r = V^T(x_i - \mu_X) \quad \text{for } i=1,2,3,\ldots,M \quad (10)$$

Y_r is the projected training vector with the $Q \times M$ dimensions.

Similarly, the projected testing data into the eigenspace can be done as follow:

$$Y_s = V^T(r_i - \mu_X) \quad \text{for } i=1,2,3,\ldots,P \quad (11)$$

where r_i is the i^{th} testing data and P is the number of testing data. Y_s is the projected testing vector with $Q \times P$ dimensions. Y_r and Y_s is used as an input features to the SVM classifier for the verification process.

Based on the three selected threshold values, three sets of principle components with sizes of 740×80, 740×117 and 740×180 are used as training data while 1480×80, 1480×117 and 1480×180 of principle component sizes are used for testing data.

4. Kernel Principal Component Analysis (KPCA) Processing

KPCA technique is the result of applying the kernel function to PCA in order to obtain the representation of PCA in a higher dimensional space. In order to perform KPCA, the training samples, x needs to be projected into the high dimensional feature space F as follows:

$$\Phi: \rightarrow F;\ x \rightarrow X \quad (12)$$

In this research, two types of kernel are applied i.e. Gaussian and polynomial kernel. The Gaussian and Polynomial kernel are given as in equation (13) and equation (14) respectively:

$$k(x_i, x_j) = \exp\left(-\frac{1}{2s^2}\left\|x_i - x_j\right\|^2\right)$$

; s is Gaussian parameter (13)

$$k(x_i, x_j) = (x_i \cdot x_j + 1)^p$$

; p is polynomial parameter (14)

Except for utilizing the kernel trick, KPCA perform the same process as PCA in projecting Y_r and Y_s. The output of the KPCA process is also a matrix in dimension of $Q \times M$ for training samples and $Q \times P$ for testing samples. The projected training and testing data for KPCA technique are set to the same dimensions as in the PCA technique.

5. Classification

In this research, the multi-class classifier is performed for the verification process. Several methods can be used to implement SVM classifier for multi-classes such as one against one, one against all and Directed Acyclic Graph Support Vector Machine (DAGSVM) method. This research uses the one against all method. Here, for N class classification, SVM requires the *N* training data to be built as a reference model, where each model is used to isolate one class from the remaining *N* classes.

Two types of SVM classifiers i.e. linear SVM and polynomial SVM have been used in this study. Linear SVM is the original optimal hyperplane algorithm that widely used as classifiers in a linearly separable case. Meanwhile, Polynomial SVM is a way to create nonlinear classifiers by applying the kernel trick.

The speaker verification system in this research uses four types of feature reduction methods i.e. MFCC, MFCC-PCA, MFCC-KPCA_Gaussian and MFCC-KPCA_Polynomial. Each type of features is evaluated using two types of SVM i.e. linear and polynomial.

6. Results and discussion

In this research, system performances will be evaluated in term for Equal Error Rate (EER) and verification time (VT). According to Kung et. al. (2005), the accuracy of biometric system is evaluated using false rejection rate (FRR) and false acceptance rate (FAR) which respectively corresponds to sensitivity and specificity. FRR which is also known as miss probability is the rejection percentage of authorized individuals while genuine acceptance rate (GAR) is the percentage of authorized individuals accepted by the verification system.

FAR which is also known as impostor pass rate is the percentage of unauthorized individuals is accepted by the verification system. FRR and FAR values can help us to determine the level of sensitivity and specificity. High FRR specify low sensitivity, while high FAR specify low specificity. A good verification system supposed to have a low FRR (high sensitivity) and low FAR (high specificity). Consequently, verification time, VT is also determined where it is the time taken by SVM classifier to verify testing data samples. This verification time is calculated to evaluate a significant time saving for verification process.

6.1. Performances based on different feature extraction techniques and different feature dimensions using linear SVM classifier

Table 1 shows the performances of speaker verification system using linear SVM based on different feature extraction techniques. The EER percentage and verification time for MFCC technique before the dimension reduction (dimension=4096) is 1.0163% and 37.4093s, respectively. The MFCC-KPCA_Gaussian technique gives the best EER performance at feature dimension of 117 with EER value equals to 0.8146% and verification time equals to 2.3385s. It is observed that the EER values for MFCC technique of smaller dimensions decrease the system performance but it can speed up the verification time.

Table 1: EER and verification time performances of different feature extraction techniques based on linear SVM classifier

Features Extraction Techniques	Features Dimension							
	D=100		D=144		D=196		D=4096	
	EER(%)	VT(s)	EER(%)	VT(s)	EER(%)	VT(s)	EER(%)	VT(s)
MFCC	5.4608	2.1413	4.5965	2.8167	3.8786	3.6361	1.0163	37.4093
Features Extraction Techniques	Features Dimensions							
	D=80		D=117		D=180			
	EER(%)	VT(s)	EER(%)	VT(s)	EER %	VT(s)		
MFCC-PCA	1.2294	2.2465	1.159	2.5117	1.0698	3.0604		
MFCC-KPCA Gaussian	1.1571	2.0806	0.8146	2.3385	0.8643	2.8530		
MFCC-KPCA Polynomial	3.6924	2.2414	2.6642	2.4330	1.8253	2.9464		

6.2. Performance based on different feature extraction techniques and different features dimension using Polynomial SVM classifier

Table 2 shows the performances of speaker verification system using polynomial SVM based on different feature extraction techniques. The EER percentage and verification time before the dimension reduction (dimension=4096) is 0.9685% and 42.5254s, respectively. It is observed that the data reduction techniques based on PCA and KPCA significantly improved the verification time and have surpassed the EER performances of MFCC method for the same category of size dimensions except for MFCC-KPCA Gaussian with 100 feature dimension.

6.3. Receiver Operating Curve based on GAR and FAR performances for selected feature dimension

Fig. 3 shows the performances of different feature extraction techniques based on selected feature dimensions according to their best EER performances using linear SVM as classifier. 100% of GAR performance for MFCC with 4096 dimensions is found at FAR equals to 40%. At the same FAR percentage, GAR performances for MFCC with 196 dimensions, MFCC-PCA with 180 dimensions, MFCC-KPCA_Gaussian with 117 dimensions and MFCC-KPCA_polynomial with 180 dimensions are 99.7%, 99.9%, 99.9% and 99.85% respectively.

Table 2: EER and verification time performances of different feature extraction techniques based on polynomial SVM classifier

Features Extraction Techniques	Features Dimension							
	D=100		D=144		D=196		D=4096	
	EER(%)	VT(s)	EER(%)	VT(s)	EER(%)	VT(s)	EER(%)	VT(s)
MFCC	4.3956	2.2326	4.0756	2.8417	3.4797	3.5331	0.9685	42.5254
Features Extraction Techniques	Features Dimensions							
	D=80		D=117		D=180			
	EER(%)	VT(s)	EER(%)	VT(s)	EER(%)	VT(s)		
MFCC-PCA	1.7258	1.5817	1.6441	3.5346	1.5888	4.5867		
MFCC-KPCA Gaussian	1.4621	1.5526	1.2885	3.0872	1.5559	4.0749		
MFCC-KPCA Polynomial	4.4125	1.6109	3.2245	3.2089	2.5375	4.1237		

Meanwhile, at FAR equals to 1%, the GAR performances are 98.99%, 90.95%, 98.92%, 99.39% and 97.43%, respectively. Table 3 shows the EER performances based on the selected features dimensions. The MFCC-KPCA_Gaussian with 117 dimensions achieves a significant improvement and outperforms the other approaches.

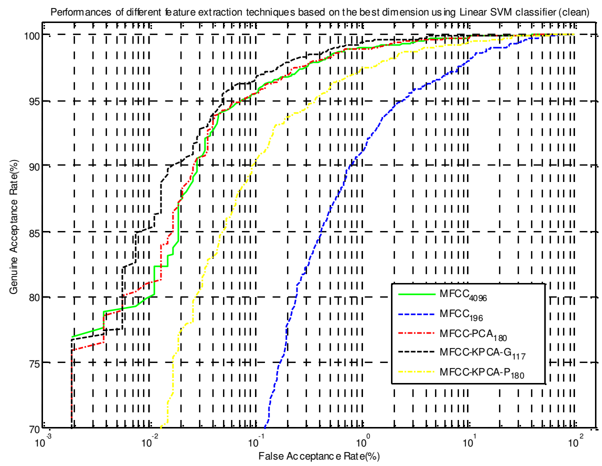

Fig. 3: Receiver operating curve (ROC) based on the selected feature dimensions for linear SVM classifier

Table 3: EER performances based on the selected feature dimensions for linear SVM classifier

Feature Extraction Technique	$MFCC_{4096}$	$MFCC_{196}$	MFCC-PCA_{180}	MFCC-KPCA-$Gauss_{117}$	MFCC-KPCA-$Poly_{180}$
EER (%)	1.0163	3.8786	1.0698	0.8146	1.8253

Subsequently, Fig. 4 shows the performances of different feature extraction techniques based on the selected features dimensions according to their best EER performances using polynomial SVM as classifier. 100% of GAR performance for MFCC with 4096 dimensions is observed at FAR equals 25.28%. At the same FAR percentage, GAR performances for MFCC with 196 dimensions, MFCC-PCA with 180 dimensions, MFCC-KPCA_Gaussian with 117 dimensions and MFCC-KPCA_polynomial with 180 dimensions are 99.26%, 99.66%, 99.93% and 99.26%, respectively. Meanwhile, at FAR equals to 1%, the GAR performances are 99.05%, 93.45%, 97.64%, 98.38% and 96.35%, respectively. Table 4 shows the EER performances based on the selected features dimensions. The MFCC with baseline dimensions of 4096 shows the best EER results. However, this is unfavourable due to the long processing time as discussed in the previous section.

7. Conclusion

As the processing time is critical in running the real time system, this study evaluated the data reduction based on principle component analysis for speech signal data. This study reveals that by executing the right method for data reduction can really improve the time taken for verification process and at the same time can maintain the system accuracy. The performances of MFCC-SVM, MFCC-PCA-SVM, MFCC-KPCA_Gaussian-SVM and MFCC-KPCA_Polynomial-SVM system have been evaluated for this purpose. Based on EER evaluation, the best performance has been observed using KPCA_Gaussian by using linear SVM as the classifier. For future research, the improvement based on speed up algorithm on kernel calculation should be considered.

Acknowledgments

The authors would like to thank the financial support provided by Universiti Sains Malaysia under Research University Grant 814161 for this project.

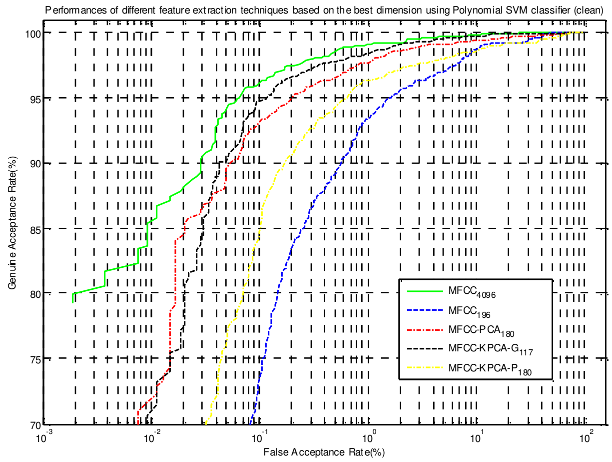

Fig. 4: Receiver operating curve (ROC) based on the selected feature dimensions for linear SVM classifier

Table 4: EER performances based on the selected feature dimensions for linear SVM classifier

Feature Extraction Technique	MFCC$_{4096}$	MFCC$_{196}$	MFCC-PCA$_{180}$	MFCC-KPCA-Gauss$_{117}$	MFCC-KPCA-Poly$_{180}$
EER (%)	**0.9685**	3.4797	1.5888	1.2885	2.5375

References

Amaro L., Heiga Z., Nankaku Y., Miyajima C., Tokuda K., and Kitamura T.; 2004. On the use of kernel PCA for feature extraction in speech recognition. IEICE Transactions on Information and Systems, 87(12), pp. 2802-2811.

Chen S.H. and Luo Y.R.; 2009. Speaker verification using MFCC and support vector machine. Proceedings of the International Multi Conference of Engineers and Computer Scientists: 18-20.

Davis S. and Mermelstein P.; 1980. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Transactions on Acoustic Speech and Signal Processing, 28(4), pp. 357-366.

Ganchev T., Fakotakis N. and Kokkinakis G.; 2005. Comparative evaluation of various MFCC implementations on the speaker verification task. Proceedings of the SPECOM 1: 191-194.

Hasan M. R., Jamil M. and Rahman M.G.R.M.S.; 2004. Speaker identification using Mel frequency cepstral coefficients. 3rd International Conference on Electrical & Computer Engineering (ICECE 2004): 565-568.

Huang S.Y., Yeh Y.R. and Eguchi S.; 2009. Robust kernel principal component analysis. Neural Computation, 21(11), pp. 3179-3213.

Ishak K.A., Samad S.A. and Hussain A.; 2006. A face detection and recognition system for intelligent vehicles. Information Technology Journal, 5(3), pp. 507-515.

Ittichaichareon C., Suksri S. and Yingthawornsuk T.; 2012. Speech recognition using MFCC. International Conference on Computer Graphics, Simulation and Modeling (ICGSM'2012): 135-138.

Lee S.M., Fang S.H., Hung J.W. and Lee L.S.; 2001. Improved MFCC feature extraction by PCA-optimized filter-bank for speech recognition. IEEE Workshop on Automatic Speech Recognition and Understanding: 49-52.

Leitner C., Pernkopf F. and Kubin G.; 2011. Kernel PCA for speech enhancement. INTERSPEECH: 1221-1224.

Mishra P. and Agrawal S.; 2012. Recognition of voice using Mel cepstral coefficient & Vector Quantization. International Journal of Engineering Research and Applications (IJERA) 2012, 2(2), pp. 933-938

Rodríguez J.M.C., de Paz F., Rocha M.P. and Riverola F.F.; 2008. Improving a leaves automatic recognition process using PCA. 2nd International Workshop on Practical Applications of Computational Biology and Bioinformatics (IWPACBB 2008): 243-251.

Sanderson C. and Paliwal K.K.; 2001. Noise compensation in a multi-modal verification system. International Conference on Acoustics, Speech, and Signal Processing: 157-160.

Turk M.A. and Pentland A.P.; 1991. Face recognition using eigenfaces. IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1991: 586-591.