BHATTACHARYYA DISTANCE BASED EMOTIONAL DISSIMILARITY MEASURE FOR EMOTION CLASSIFICATION

Tin Lay Nwe, Nguyen Trung Hieu, Dilip Kumar Limbu

Human Language Technology Department, Institute for Infocomm Research (I2R), A*STAR, Singapore 138632

{tlnma,thnguyen,dklimbu}@i2r.a-star.edu.sg

ABSTRACT

Speech is one of the most important signals that can be used to detect human emotions. When speech is modulated by different emotions, spectral distribution of speech is changed accordingly. A Gaussian Mixture Model(GMM) can model the changes in spectral distributions effectively. A GMM-supervector characterizes the spectral distribution of an emotion utterance by the GMM parameters such as the mean vectors and covariance matrices. In this paper, we propose to use the GMM-supervectors that characterize the emotional spectral dissimilarity measure for emotion classification. We employ the GMM-SVM kernel with Bhattacharyya based GMM distance to obtain dissimilarity measure. Beside the first-order statistics of mean, we consider dissimilarity measure using second-order statistics of covariance which describe the shape of the distribution. Experiments are conducted using SVM classifier to classify emotions of anger, happiness, neutral and sadness. We achieve average accuracy of 78.14% for speaker independent emotion classification.

***Index Terms*—** Emotion classification, emotional dissimilarity measure, Gaussian Mixture Model (GMM), supervector, Support Vector Machine (SVM)

1. INTRODUCTION

Speech based emotion classification is an important research topic in the area of Human Computer Interaction (HCI) and has a wide range of applications. An example of such application is car board system where information of the mental state of the driver may be provided to initiate his/her safety. Another example is call center application in which emotion classification system can be used to detect customer's satisfaction.

Emotion classification system includes two modules. The first module is front-end feature extraction and the second module is back-end classifier. Several back-end classifiers are reported in the literature. These include support vector machine (SVM) [1], [2], [3], [4], Gaussian Mixture Model (GMM) [5], Support Vector Regression (SVR) [6] and Neural Network [7]. SVM is a novel type of learning machine, which is an approximate implementation of the method of structural risk minimization. SVM has shown to provide a better generalization performance in solving various classification problems than traditional techniques [7]. In this paper, we use SVM as the back-end classifier.

Regarding the front-end feature extraction component, a number of features have been explored for SVM classifier in different studies. The studies [2], [3], employ prosodic features based on statistics of pitch, energy and duration, higher order formants for emotion classification. In addition to prosodic features, the study in [7] investigates zero crossing rate, spectrum centroid, spectrum cut-off frequency, correlation density and Mel-frequency energy. In [4], features such as Linear Prediction Cepstral Coefficients (LPCC), Mel-Frequency Cepstral Coefficients (MFCC), Log Frequency Power Coefficients (LFPC), Perceptual Linear Prediction (PLP) are used to classify emotions. The study [1] proposes Modulation Spectral Feature (MSF) which is obtained using an auditory filterbank and a modulation filterbank for speech analysis. MSF feature is able to capture both acoustic frequency and temporal modulation frequency components. The features explored in the above studies are related to the acoustic characteristics of frequency, energy, and spectral intensity. These are basic characteristics of acoustic signals. By extracting acoustic features with these basic characteristics, the emotion contained in a speech signal can be recognized.

In fact, acoustic characteristics of individual emotions are different. Pitch contour of anger emotion has irregular up and down inflection [8] while that of happiness has descending line [9]. Furthermore, mean pitch values of anger and happiness are high [8] while that of sadness is below normal mean [10]. And, spectral energy is concentrated in high frequency regions in anger emotion, but in sad emotion, the spectral energy is concentrated in low frequency regions [11]. Based on these investigations, we propose features that reflect emotional spectral dissimilarity measure which will be referred to as Emotional Dissimilarity (ED) measure for emotion classification.

The Bhattacharyya distance was first introduced in [12]. It gives better results than the KL divergence in several applications of classical statistics [13]. In [14], Bhattacharyya distance based GMM-supervector is successfully employed to the task of speaker recognition. In this paper, we propose to use Bhattacharyya distance based GMM-supervectors [14] for emotion classification and present the usefulness of Bhattacharyya distance to measure Emotional Dissimilarity (ED). ED measure based on first-order statistics of mean gives the major characteristics of the probabilistic distance. Beside the the first-order statistics, we also consider the ED measure on spectral shape which is obtained by using second-order statistics of covariance. The advantage of integrating the second-order statistics to the GMM-supervectors with first-order statistics is that GMM-supervectors become more descriptive to characterize an unknown emotion.

Our work is related to the prior study [15]. In the study [15], GMM-supervector based SVM with spectral features is used for emotion recognition. This study employs KL divergence kernel [16] to construct the GMM-supervector and considers only the first-order statistics of mean. In our study, we use both first and second-order statistics in GMM-supervector formulation. Furthermore, we em-

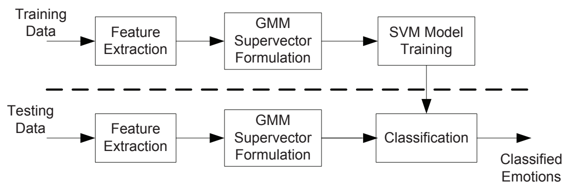


Fig. 1: *System Block Diagram*

ploy Bhattacharyya distance based kernel which is proven to perform better than KL based kernel for the case of speaker recognition [14]. We compare the performance of the Bhattacharyya distance based GMM-supervector with those of the KL based GMM-supervector and baseline generalized GMM-supervector formulated without applying kernel.

The block diagram of the emotion classification system is shown in Fig. 1. We extract acoustic features from each emotional utterance. Then, we formulate GMM-supervectors to use as features in SVM classifier. We consider to classify emotions namely anger, happiness, neutral and sadness.

The rest of the paper is organized as follows. In section 2, we present the emotion corpus used in our experiments. In section 3, we describe the GMM-supervector formulation using GMM-SVM kernels. In section 4, we present our experiments and results. Finally, we conclude our study in section 5.

2. EMOTION CORPUS

We collected emotional speech data for our experiments. Database includes English emotion utterances portrayed by an actress and five housewives. The database is referred to as Emotions of an Actress And Housewives (EAH). A total of seven emotions: anger, dislike, fear, happiness, neutral, sadness and surprise are included. A total of 400 emotionally neutral sentences are prepared. An actress repeated each of 400 sentences with each of seven emotions. Hence, a total of 400 $\times$ 7 = 2800 utterances are obtained from an actress. And, each housewife repeated each of 100 sentences with each of seven emotions. Hence, a total of 100 $\times$ 7 $\times$ 5 = 3500 utterances are obtained from five housewives. Average length of the utterances is 4.8 seconds with standard deviation of 1.2. There are 900 utterances for each of seven emotions. Utterances are sampled at 16kHz and 16 bit rate. We use the subset of EAH database including emotions of anger, happiness, neutral and sadness in our experiments.

3. GMM-SUPERVECTOR FORMULATION

The GMM-supervector can be considered of as a mapping between an utterance and a high-dimensional vector through a kernel [16]. Kernels are important components for SVM learning. It is a method of using a linear classifier to solve a non-linear problem by nonlinearly mapping the original observations into a higher-dimensional space, where a linear classifier is subsequently used. This makes linear classification in the new feature [17]. In the following subsections, we discuss the GMM-supervector formulation using GMM-SVM kernels based on KullbackLeibler (KL) divergence and Bhattacharyya based GMM distance. We also present the GMM-supervector formulation without employing kernel function. This formulation is referred to as Generalized GMM-supervector.

3.1. Generalized GMM-supervector

The density function of a GMM is defined as in equation (1).

$$p\left(x\right) = \sum_{i=1}^{M} \omega_i f\left(x|m_i, \Sigma_i\right) \tag{1}$$

where $f\left(.\right)$ denotes the Gaussian density function. And, m_i, Σ_i and ω_i are the mean, covariance matrix and weight of i^{th} Gaussian component, respectively. M is number of Gaussian mixtures. And, x is a D-dimensional acoustic feature vector. We formulate the Generalized GMM-supervector by stacking mean vectors of the GMM.

3.2. GMM-supervector with KL based kernel

The Kullback-Leibler (KL) divergence, also known as mutual information, relative entropy or, simply, information divergence, is a classic information gain measure of the asymmetric difference between two distributions, a and b, i.e. it measures the divergence from one probability distribution to another as in equation (2) [14]. Hence, KL divergence can also be used to measure ED.

$$\begin{aligned}\psi^{KL}\left(f\left(m_i^a, \Sigma_i^a\right) \| f\left(m_i^b, \Sigma_i^b\right)\right) = \\ \frac{1}{2}\left\{\ln\left(\frac{|\Sigma_i^b|}{|\Sigma_i^a|}\right) + \text{tr}\left(\left(\Sigma_i^b\right)^{-1}\Sigma_i^a\right)\right. \\ \left. + \left(m_i^b - m_i^a\right)^T \left(\Sigma_i^b\right)^{-1}\left(m_i^b - m_i^a\right) - D\right\}\end{aligned} \tag{2}$$

KL divergence is neither positive definite nor symmetric. To satisfy Mercer's condition [16], symmetrized version of the KL divergence is used to formulate KL based kernel [16], [18]. The kernel includes a mean vector term and a covariance term as in equation (3).

$$\begin{aligned}K^{KL}\left(X_a, X_b\right) = \\ \sum_{i=1}^{M}\left(\sqrt{\omega_i}\left(\Sigma_i^u\right)^{-1/2} m_i^a\right)^T \left(\sqrt{\omega_i}\left(\Sigma_i^u\right)^{-1/2} m_i^b\right) \\ + \sum_{i=1}^{M}\frac{\omega_i}{2}\text{tr}\left(\Sigma_i^a\left(\Sigma_i^u\right)^{-2}\Sigma_i^b\right)\end{aligned} \tag{3}$$

where tr is the trace of the matrix, Σ_i^λ is the adapted covariance matrix, and Σ_i^u is the covariance matrix of the Universal Background Model (UBM). X_a and X_b denote acoustic feature vector sequences of utterances a and b respectively. Based on this kernel, the i^{th} subvector of the GMM-supervector is formulated as in equation (4) [14].

$$g^{KL}(m_i, \Sigma_i) = \begin{bmatrix} \sqrt{\omega_i}\left(\Sigma_i^u\right)^{-1/2} m_i^\lambda \\ \sqrt{\frac{\omega_i}{2}}\text{diag}\left(\Sigma_i^\lambda\left(\Sigma_i^u\right)^{-1}\right)\end{bmatrix} \tag{4}$$

GMM-supervector with KL based kernel is formulated by stacking all i^{th} subvectors of equation (4).

3.3. GMM-supervector with Bhattacharyya based kernel

Bhattacharyya distance is a separability measure between two Gaussian distributions [19]. The Bhattacharyya distance between the two probability distributions is defined as in equation (5) [14]. In

$$\Psi^{Bhat}(p_a \| p_b) \approx \frac{1}{8}\sum_{i=1}^{M}\left\{\left[\left(\frac{\Sigma_i^a + \Sigma_i^u}{2}\right)^{-1/2}(m_i^a - m_i^u)\right]^T\left[\left(\frac{\Sigma_i^b + \Sigma_i^u}{2}\right)^{-1/2}\left(m_i^b - m_i^u\right)\right]\right\}$$
$$+\frac{1}{2}\sum_{i=1}^{M}\mathrm{tr}\left[\left(\frac{\Sigma_i^a + \Sigma_i^u}{2}\right)^{1/2}(\Sigma_i^a)^{-1/2}\left(\frac{\Sigma_i^b + \Sigma_i^u}{2}\right)^{1/2}\left(\Sigma_i^b\right)^{-1/2}\right]$$
$$+\frac{1}{2}\sum_{i=1}^{M}\left\{\frac{\omega_i^u}{\omega_i^a}\frac{\omega_i^u}{\omega_i^b}\right\} - \sum_{i=1}^{M}\ln\{\omega_i^u\} - M \tag{5}$$

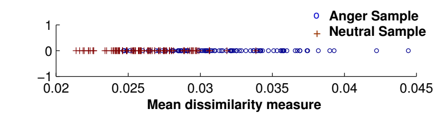
equation (5) p_a and p_b are the probabilistic models, GMM_a and GMM_b, respectively.

The first term of equation (5) gives the class separability due to the difference between class means, while the second term gives the class separability due to the variance between class covariance. Based on the first two terms, Bhattacharyya distance based kernel is formulated as in equation (6) [14]. Based on this kernel, the i^{th} subvector of the GMM-supervector is formulated as in equation (7)[14]. GMM-supervector with Bhattacharyya based kernel is obtained by stacking all i^{th} subvectors of equation (7).

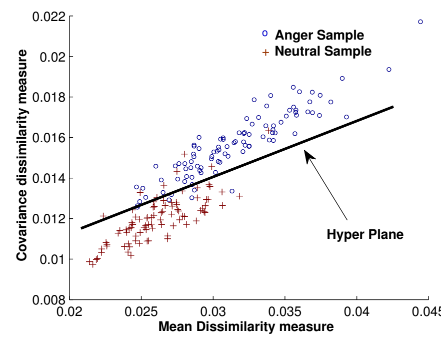
$$g^{Bhat}(m_i, \Sigma_i) = \begin{bmatrix} \left(\frac{\Sigma_i^\lambda + \Sigma_i^u}{2}\right)^{-1/2}\left(m_i^\lambda - m_i^u\right) \\ \mathrm{diag}\left(\left(\frac{\Sigma_i^\lambda + \Sigma_i^u}{2}\right)^{1/2}\left(\Sigma_i^\lambda\right)^{-1/2}\right) \end{bmatrix} \tag{7}$$

If we look at equation (7), the first term reflects the dissimilarity between mean of an emotion utterance and that of a UBM. This mean statistical dissimilarity gives the major characteristics of the probabilistic distance. And, this term represents the ED measure between an emotion utterance from a reference UBM. If a reference UBM is trained using neutral utterances, this term is to measure the ED of an emotion utterance from a neutral UBM. Besides the first-order statistics of mean, the second-order statistics of covariance matrices describing the shape of the spectral distribution is also useful to measure ED. If we look at the second term of equation (7), it represents the ratio between covariance of a UBM and that of an emotion utterance. In other words, this second term describes ED measure in terms of spectral shape.

We draw the scatter plot as illustrated in Figure 2 to compare the capability of the two GMM-formulations: 1) using only first-order term and 2) using both first and second-order terms in classifying neutral and anger emotions. In the figure, mean dissimilarity measure is calculated by taking average over absolute values of the first term in equation (7) for each utterance. And, covariance dissimilarity measure is the standard deviation of the second term. Each marker ('o' or '+') represents an emotion utterance. As we can see in the Figure 2(a), the boundary between the two classes is not clear and is errorful when we use only first-order term. However, boundary between the two emotion classes becomes clearer when we use both terms as illustrated in Figure 2(b).



(a) Using only mean dissimilarity measure



(b) Using both mean and covariance dissimilarity measure

Fig. 2: Scatter plot of distinguishing anger and neutral samples.

4. EXPERIMENTS AND RESULTS

We conduct several experiments to investigate the effectiveness of proposed approaches. Firstly, we compare the effectiveness of employing ED measure in formulating GMM-supervectors. Secondly, we observe the effect of integrating second-order statistics of covariance term in GMM-supervector formulation. Finally, we compare the performance of standard GMM classifier with that of SVM classifier with Bhattacharyya distance based kernel.

Each emotion utterance is divided into 20ms frames with 10ms overlapping. Each frame is multiplied by a Hamming window to minimize signal discontinuities at the end of each frame. From each frame, we extract Mel-Frequency Cepstral Coefficients (MFCC), Linear Predictive Cepstral Coefficient (LPCC) and Perceptual Linear Prediction Coefficients (PLPC) features. Each feature has 12 coefficients and their first derivatives. We form a feature vector for each frame by concatenating all three MFCC, LPCC and PLPC features. As each feature has a total of 24 coefficients, a feature vector of a frame has 72 coefficients.

To formulate GMM-supervectors, we extract the features from each emotion utterance. Then, we use maximum a posteriori (MAP) criterion [20] to adapt the GMM from a Universal Background Model(UBM) for each utterance. We train UBM via EM algorithm [21] using 64 mixtures. We adapt the mean and covariance only. Once we have an adapted GMM model, we formulate GMM-supervectors using the techniques mentioned in sections 3. We use National Institute of Standards and Technology (NIST) speaker recognition evaluation (SRE) 2001 dataset [22] to train UBM model in all experiments. We use SVMTorch [23] for the training and testing of SVM. We employ the target model against anti-model

$$K^{Bhat}(X_a, X_b) = \sum_{i=1}^{M} \left\{ \left[\frac{1}{2} \left(\frac{\Sigma_i^a + \Sigma_i^u}{2} \right)^{-1/2} (m_i^a - m_i^u) \right]^T \left[\frac{1}{2} \left(\frac{\Sigma_i^b + \Sigma_i^u}{2} \right)^{-1/2} \left(m_i^b - m_i^u \right) \right] \right\}$$
$$+ \sum_{i=1}^{M} \mathrm{tr} \left[\left(\frac{\Sigma_i^a + \Sigma_i^u}{2} \right)^{1/2} (\Sigma_i^a)^{-1/2} \left(\frac{\Sigma_i^b + \Sigma_i^u}{2} \right)^{1/2} \left(\Sigma_i^b \right)^{-1/2} \right] \quad (6)$$

strategy for multi-class classification.

We use the subset of EAH database including anger, happiness, neutral and sadness emotions to perform speaker independent emotion classification. EAH database includes 6 speakers. In all experiments, we perform 6 folds cross validation. In each fold, emotion samples of a speaker are used as test samples and those of the remaining five speakers are used as training samples. The estimated classification accuracy is the mean the accuracy over 6 folds. In the following sections we presents the experiments conducted and results obtained.

4.1. Effect of GMM-supervector with ED measure

To investigate ED measure, we formulate the GMM-supervector using the first term of equation (7) with the Bhattacharyya distance kernel. We also formulate the GMM-supervector using the first term of equation (4) with KL based kernel. The first terms in the above equations use only first-order statistics. And, we formulate the generalized GMM-supervectors by concatenating mean terms of an adapted GMM. We perform emotion classification experiments using these 3 different GMM-supervectors and the average accuracies achieved are shown in the second row of Table 1.

Table 1: *Speaker independent average emotion classification accuracies (%)*

Terms in GMM-supervector	GEN	KL	$Bhat$
first-order	59.13	66.36	76.48
first-order + second-order	60.56	68.58	78.14

Both Bhattacharyya and KL based kernels are attributed to the approximation of dissimilarity measure between two distributions [14], [15]. The results show that GMM-supervectors that employ ED measure performs better than baseline generalized GMM-supervector. Of the two ED based GMM-supervectors, Bhattacharyya kernel based one is better than KL based one. If we compare the kernels mentioned in equations (3) and (6) , the latter one with Bhattacharyya kernel includes the term such as $(m_i^a - m_i^u)$ which gives an absolute distance measure between an emotion utterance and a reference UBM. This characteristic fits very well with the measuring ED for the task of emotion classification. KL based kernel does not involve such term that measures the absolute distance.

4.2. Effect of integrating second-order term in GMM-supervector

We concatenate the respective second-order terms in 3 different GMM-supervectors of the subsections 4.1. Please refer to equations (7) and (4) for second-order terms of Bhattacharyya and KL based kernels respectively. As for generalized GMM-supervector, we stack the first order and second-order terms together. The average emotion classification accuracies achieved using 3 different GMM-supervectors formulations are shown in 3^{rd} row of Table 1.

As we can see in Table 1, second-order term helps to improve the emotion classification accuracies in all GMM-supervector formulations. Improvement in terms of absolute accuracies are 1.43%, 2.22% and 1.66% for Generalized, KL kernel based and Bhattacharyya kernel based GMM-supervectors respectively. Hence, second-order term of covariance matrix describing shape of the spectral distribution contributes to performance improvement. The system using GMM-supervector with Bhattacharyya based kernel performs the best in all experiments and achieves the accuracy of 78.14%.

4.3. GMM vs. SVM with Bhattacharyya kernel based GMM-Supervector

We compare the SVM system using Bhattacharyya kernel based GMM-supervector with standard GMM system. In the standard GMM system, emotion models were trained via EM algorithm [21]. Each GMM has 64 Gaussian components. A maximum likelihood Bayes classifier is used for decision. An average emotion classification accuracy obtained using standard GMM classifier is presented in first column of Table 2. We repeat the accuracies of Bhattacharyya kernel based SVM systems from Table 1 in 2^{nd} and 3^{rd} columns respectively for ease of comparison.

Table 2: *Speaker independent average emotion classification accuracies (%) for GMM and SVM systems (In 1^{st} row, subscript f: first-order and subscript fs: first and second-order)*

GMM	$Bhat_f$	$Bhat_{fs}$
75.5	76.48	78.14

We can see that Bhattacharyya kernel based SVM outperforms the standard GMM for speech emotion recognition. The accuracies of SVM systems are about 1% higher for the system using only first-order term and 2.64% higher for the system using both first and second-order terms.

5. CONCLUSIONS

We have presented an approach to employ Emotional Dissimilarity (ED) measure in GMM-supervector formulation for SVM classifier to classify emotions. We investigate both first-order statistics of mean and second-order statistics of covariance terms in GMM-SVM kernel based supervectors. We employ Bhattacharyya distance based kernel to characterize Emotional Dissimilarity (ED) measure. We found that Bhattacharyya distance based kernel is better than KL based kernel to characterize ED. Furthermore, we found that ED measure using second-order statistics of covariance contributes to performance improvement. Finally, the experimental results show that Bhattacharyya kernel based SVM system performs better than standard GMM system.

6. REFERENCES

[1] Wu, S., Falk, T., Chan, W., ”Automatic Speech Emotion Recognition Using Modulation Spectral Features”. Speech Communication 53(5):768-785, 2011.

[2] Schuller B., Reiter S., Mueller R., Al-Hames M., Lang M., Rigoll G. ”Speaker-Independent Speech Emotion Recognition by Ensemble Classification”. Proc. ICME 2005, Amsterdam, Netherlands, 2005.

[3] You. M, Chen. C, Bu. J, Liu. J, Tao. J, ”Emotional Speech Analysis on Nonlinear Manifold”. In: 18th International Conference on Pattern Recognition (ICPR), Hong Kong, pp. 91-94, 2006.

[4] T.L. Pao, Y.T Chen, J.H. Yeh and P.J. Li, ”Mandarin emotional speech recognition based on SVM and NN”, 18th International Conference on Pattern Recognion, vol.1, pp.1096-1100, 2006.

[5] Kockmann, M., Burget, L., and Cernocky, J., ”Application of Speaker- and Language Identification State-of-the-Art Techniques for Emotion Recognition”, In: Speech Communication, Vol. 53, No. 9, Amsterdam, NL, p. 1172-1185, ISSN 0167-6393, 2011.

[6] Grimm, M., Kroschel, K., Narayanan, S., ”Support vector regression for automatic recognition of spontaneous emotions in speech”. In: ICASSP, vol. 4, pp. 1085-1088 (2007).

[7] Chen, L., Mao, X., Xue, Y-L., and Cheng, L. L., ”Speech Emotion Recognition: Features and Classification Models”. Digital Signal Processing 22(6): 1154-1160, 2012.

[8] Davitz, J.R., ”The Communication of Emotional Meaning”. McGraw-Hill, New York, 1964.

[9] Fonagy, I., and Magdics, K., ”Emotional patterns in intonation and music”. Z. Phonet. Sprachwiss. Kommunikationsforsch 16, 293-326, 1963.

[10] Murray, I., and Arnott, J., ”Toward the Simulation of Emotion in Synthetic Speech: A Review of the Literature on Human Vocal Emotion”, J. of Acoustical Society of America, pp. 1097-1108, 1993.

[11] Williams, C., and Stevens, K., ”Vocal Correlates of Emotional States”. Grune and Stratton, 1981, ch. Speech Evaluation in Psychiatry, pp. 189-220, 1981.

[12] Bhattacharyya, A., ”On A Measure of Divergence Between Two Statistical Populations Defined by Their Probability Distributions”, Bull. Calcutta Math. Soc., vol. 35, pp. 99-109, 1943.

[13] Kailath, T., ”The Divergence and Bhattacharyya Distance Measures in Signal Selection”, IEEE Trans. Commun. Technol., vol. COM-15, no. 1, pp. 52-60, Feb. 1967.

[14] You, C., Lee, K.A., and Li, H., ”GMM-SVM Kernel With A Bhattacharyya-Based Distance For Speaker Recognition” IEEE Trans. Audio, Speech, and Language Processing, vol. 18, no. 6, pp. 1300 - 1312, 2010.

[15] Hu, H., Xu, M. and Wu, W., ”GMM Supervector Based SVM With Spectral Features For Speech Emotion Recognition”, in ICASSP, Honolulu, Hawai, USA, April 2007, pp. 413-416.

[16] Campbell, W.M., Sturim, D.E., Reynolds, D.A., and Solomonoff, A., ”SVM based Speaker Verification Using A GMM Supervector Kernel and NAP Variability Compensation”, in Proc. of ICASSP, pp. 97-100, France, 2006.

[17] Vaiciukynas, E., Gelzinis, A., Bacauskiene, M., Verikas, A., and Vegiene, A., ”Exploring Kernels in SVM-Based Classification of Larynx Pathology from Human Voice”, Proceedings of the 5th International Conference on Electrical and Control Technologies ECT-2010, May 6-7, 2010, Kaunas, Lithuania

[18] Campbell, W. M., ”A Covariance Kernel For SVM Language Recognition”, in Proc. Int. Conf. Acoust. Speech Signal Process., 2008, pp. 4141-4144.

[19] Mak, B., and Barnard, E., ”Phone Clustering Using The Bhattacharyya Distance”. ICSLP 1996

[20] Gauvain, J. L., and Lee, C. H., ”Maximum A Posteriori Estimation For Multivariate Gaussian Mixture Observations Of Markov Chains”, IEEE Trans. Speech Audio Process., vol. 2, no. 2, pp. 291-298, 1994.

[21] Dempster, A.P., Laird, N.M., Rubin, D.B., ”Maximum Likelihood from Incomplete Data via the EM Algorithm”, Journal of the Royal Statistical Society, Series B, 39: 1-38, 1977.

[22] Przybocki, M., and Martin, A., ”The NIST Year 2001 Speaker Recognition Evaluation Plan”, http://www.itl.nist.gov/iad/mig/tests/spk/2001/2001-spkrec-evalplan-v05.9.pdf (2001)

[23] Collobert, R., and Bengio, S., ”SVMTorch: Support Vector Machines for Large-Scale Regression Problems”, J. Mach. Learn. Res., vol. 1, pp. 143-160, 2001.