



# Bhattacharyya Distance based Emotional Dissimilarity Measure in Multi-Dimensional Space 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@i2r.a-star.edu.sg, thnguyen@i2r.a-star.edu.sg, dklimbu@i2r.a-star.edu.sg

## Abstract

Emotion classification is essential for understanding human interactions and hence is an important module in Human-Computer Interaction (HCI) systems. Well-performed emotion classification systems have potential to integrate into the HCI systems to provide additional user state details. This paper presents an emotion classification system that employs Emotional Dissimilarity (ED) measure. Instead of measuring ED in Single Dimensional (SD) space, we propose an approach to measure ED in Multi-Dimensional (MD) space. Proposed approach measures ED between emotions using GMM-SVM kernel with Bhattacharyya based GMM distance. In addition, we also formulate ED measure using GMM-SVM kernel with Kullback Leibler (KL) based GMM distance. We observe the effectiveness of employing ED measure in MD space over that in SD space using different kernels. Experiments were conducted using SVM classifier to classify emotions of anger, happiness, neutral and sadness. We achieve average accuracy of 81.25% for speaker independent emotion classification.

**Index Terms:** emotion classification, emotional dissimilarity measure, supervector

## 1. Introduction

Emotion is an important paralinguistic information in human communication. Emotion directs non-linguistic social signals such as body language, facial expression, etc., to express wants, needs and desires [1]. There are many applications for emotion classification system in the fields of healthcare, services, telecommunication, etc.,. In the healthcare field, emotion classification system can help clinicians to access psychological disorders online arising from emotional difficulties [1]. As for services, emotion classification system can be used to detect customer's satisfaction. An example is a call center application. For telecommunication, emotion classification system can be used to route 911 emergency call services for high priority emergency calls.

An emotion classification system has two major components: 1) feature extraction and 2) classifier. Several classifiers are reported in the literature. These include support vector machine (SVM) [2], [3], [4], [5], [6], Gaussian Mixture Models (GMM) [7], Support Vector Regression (SVR) [8], weighted linear regression [9] and Neural Network [10]. In these classifiers, SVM has shown to provide a better generalization performance in solving various classification problems than traditional techniques [10]. In this paper, we use SVM to classify emotions.

For feature extraction component, a number of acoustic features have been explored for SVM classifier in different studies.

In the studies [3], [4], prosodic features based on statistics of pitch, energy, duration and higher order formants are studied. In addition to prosodic features, the study in [10] investigates zero crossing rate, spectrum centroid, spectrum cut-off frequency, correlation density and Mel-frequency energy. In [5], 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. A number of 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.

In fact, different emotions have different acoustic characteristics. When speech is modulated by different emotions, spectral distribution of speech is changed accordingly. Hence, features that characterize emotional spectral dissimilarity measure which is referred to as Emotional Dissimilarity (ED) measure can be used for emotion classification. In our earlier work [11], we employed GMM supervectors that characterize ED for emotion classification. We employ Bhattacharyya based GMM distance to measure ED. In that study, ED measurement was in Single Dimension (SD). In this paper, we extend our earlier work [11] to propose an approach for measuring ED in Multi-Dimensions (MD). As in [11], we formulate GMM supervectors using Bhattacharyya based kernel and KL based kernel. In addition, we also formulate generalized GMM-supervectors without applying kernel. We observe the effectiveness of employing ED measure in MD over that in SD using 3 different GMM supervectors.

Our emotion classification system comprises 3 modules. In the first module, we extract acoustic features from each emotion utterance. Then, we formulate GMM-supervectors in the second module. Finally, the third module is 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 review the GMM supervector formulation methods in SD. In section 3, we discuss about GMM-supervector formulation in MD. In section 4, we present our experiments and results. Finally, we conclude our study in section 5.

## 2. Review on GMM-supervector formulation in Single Dimension (SD)

SVM is a linear classifier. But SVM classifier can solve non-linear problems by using kernels. In the following subsections, we review kernel based GMM-supervector formulation methods used in our earlier work [11]. The methods use GMM-SVM kernels based on KullbackLeibler (KL) divergence and Bhat-

tacharyya based GMM distance. Furthermore, we will review Generalized GMM-supervector formulation which does not apply kernel function.

### 2.1. Generalized GMM-supervector

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

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

where  $f(\cdot)$  denotes the Gaussian density function. And,  $m_i, \Sigma_i$  and  $\omega_i$  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 [12].

### 2.2. 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, i.e. it measures the divergence from one probability distribution to another [13]. Hence, KL divergence can be used to measure ED. The  $i^{th}$  subvector of GMM supervector with KL based kernel is presented in equation (2) [12]. In equation (2), the first term is a mean vector term and the second term is a covariance term.

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

where  $m_i^\lambda$  and  $\Sigma_i^\lambda$  is the adapted mean and covariance matrix. And,  $\Sigma_i^u$  is the covariance matrix of a UBM. GMM-supervector with KL based kernel is formulated by stacking all  $i^{th}$  subvectors of equation (2).

### 2.3. GMM-supervector with Bhattacharyya based kernel

Bhattacharyya distance is a separability measure between two Gaussian distributions [14]. The  $i^{th}$  subvector of the GMM-supervector based on Bhattacharyya distance is formulated as in equation (3)[12]. GMM-supervector with Bhattacharyya based kernel is obtained by stacking all  $i^{th}$  subvectors of equation (3). If we look at equation (3), 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. And, the second term of equation (3) gives the class separability due to the variance between class covariance.

$$g^{Bhat}(m_i, \Sigma_i) = \begin{bmatrix} \left( \frac{\Sigma_i^\lambda + \Sigma_i^u}{2} \right)^{-1/2} (m_i^\lambda - m_i^u) \\ \text{diag} \left( \left( \frac{\Sigma_i^\lambda + \Sigma_i^u}{2} \right)^{1/2} (\Sigma_i^\lambda)^{-1/2} \right) \end{bmatrix} \quad (3)$$

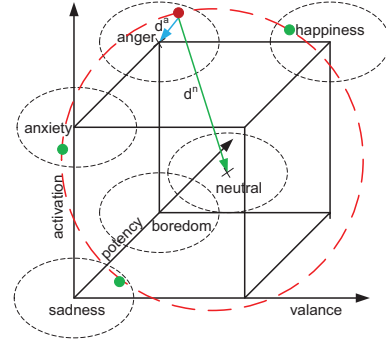


Figure 1: Illustrating effectiveness of using ED measure in Multi-Dimension (MD) to identify an unknown emotion

The above GMM supervector formulation methods involve measuring ED in Single Dimension (SD) as it measures the distance from only one reference point, the neutral UBM. In the following section, we will present the GMM-supervector formulation using ED measure in Multi-Dimensions (MD).

## 3. GMM-supervector formulation in Multi-Dimensions (MD)

Emotions are complex constructs with fuzzy boundaries and with substantial individual variations in expression and experience. To better model emotions, the study [15] uses 3 Dimensional (3D) emotion space: activation (arousal), potency (power), and valance (pleasure, evaluation) as in Figure 1. Emotion space includes six basic emotions: anger, happiness, sadness, anxiety, boredom and neutral. In Figure 1, we can see an anger emotion on dotted big circle and it is  $d^n$  distance away from neutral. In fact, all other points on the circle are also  $d^n$  distance away from neutral and they belong to different emotions. Here, distance or ED measure,  $d^n$ , is made from only one reference point and we refer this measurement as ED measure in single dimension. If different emotions have same distances from a neutral point, we can not identify emotions using ED measure in single dimension. Hence, we propose ED measure in Multi-Dimension (MD) in which we use additional distance measurements. As an example, we add one more distance measurement  $d^a$  from anger reference point in Figure 1. As this approach involves measuring distances from multiple reference points, it is refers to as ED measure in Multi-Dimension (MD). When we use ED measure in MD, we can identify anger emotion correctly as shown in Figure 1. Based on this investigation, Emotional Dissimilarity (ED) measure in Multi-Dimension (MD) is useful, especially to the cases in which different emotions have equal distances from neutral.

We formulate GMM supervectors that employ ED measure in MD. We train UBMs for each of four emotions: anger, happiness, neutral and sadness to obtain multiple reference points. Equation (4) presents  $i^{th}$  subvector of Bhattacharyya distance based GMM-supervector,  $g_a^{Bhat}$ , in anger dimension.

$$g_a^{Bhat}(m_i, \Sigma_i) = \begin{bmatrix} \left( \frac{\Sigma_i^\lambda + \Sigma_i^{u_a}}{2} \right)^{-1/2} (m_i^\lambda - m_i^{u_a}) \\ \text{diag} \left( \left( \frac{\Sigma_i^\lambda + \Sigma_i^{u_a}}{2} \right)^{1/2} (\Sigma_i^\lambda)^{-1/2} \right) \end{bmatrix} \quad (4)$$

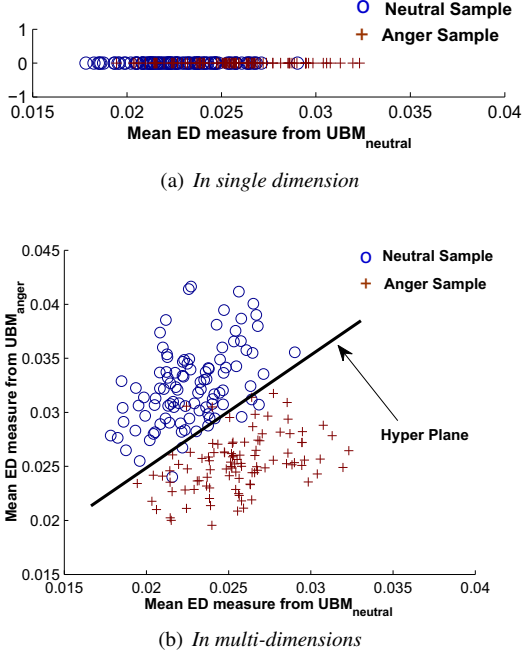


Figure 2: Scatter plot of distinguishing anger and neutral samples.

where  $u_a$  is an anger UBM. We obtain GMM-supervector,  $g_a^{Bhat}(m, \Sigma)$ , using anger UBM,  $u_a$ , by concatenating all  $i^{th}$  subvectors  $g_a^{Bhat}(m_i, \Sigma_i)$  of equation (4). Similarly, we obtain GMM-supervectors,  $g_h^{Bhat}(m, \Sigma)$ ,  $g_n^{Bhat}(m, \Sigma)$  and  $g_s^{Bhat}(m, \Sigma)$  using happiness UBM ( $u_h$ ), neutral UBM ( $u_n$ ), and sadness UBM ( $u_s$ ) respectively. Finally, Bhattacharyya distance based GMM-supervector with ED measure in MD,  $g_a^{Bhat}(m, \Sigma)$ , is obtained by concatenating GMM-supervectors  $g_a^{Bhat}(m, \Sigma)$ ,  $g_h^{Bhat}(m, \Sigma)$ ,  $g_n^{Bhat}(m, \Sigma)$  and  $g_s^{Bhat}(m, \Sigma)$  of all four emotions.

To investigate the effectiveness of employing ED measure in MD in GMM-supervector formulation, we formulate two other GMM-supervectors. One is GMM-supervector using KL based kernel,  $g_{MD}^{KL}(m, \Sigma)$  and another one is generalized GMM-supervectors,  $g_{MD}^{Gen}(m, \Sigma)$ . The process to formulate these two GMM supervectors is the same as in  $g_{MD}^{Bhat}(m, \Sigma)$  supervector formulation process as mentioned above.

We compare the effectiveness of employing ED measure in MD over that in SD to classify emotions. We use Bhattacharyya distance based GMM supervector in this comparison. Classifying neutral and anger emotions using ED measure in SD is illustrated in Figure 2(a). In the figure, mean ED measure is calculated by taking average over absolute values of the first term in equation (3) for each utterance. UBM is trained using neutral speech samples. Each marker ('o' or '+') represents an emotion utterance. As we can see in the figure, we can not properly locate a boundary to divide the two emotions: anger and neutral. Then, we investigate the ED measure in two dimensions to see if we can have clear boundary between the two classes. We train one more UBM using anger emotion samples to measure ED in one more dimension. This is illustrated in Figure 2(b). In this figure, X axis represents mean EDs of emotion samples from neutral UBM. And, Y axis represents

mean EDs of emotion samples from anger UBM. As we can see in the Figure 2(b), boundary between the two emotion classes becomes clearer when we use dissimilarity information in two dimensions.

We compute GMM-supervectors in MD based on the approaches mentioned above for each emotion utterance and perform emotion classification using SVM classifier. Experiments conducted and results obtained are given in the following section.

## 4. Experiments and Results

We conduct several experiments to investigate the effectiveness of proposed approach. Firstly, we repeat the experiments of our earlier work [11]. In this experiment, we perform emotion classification using GMM-supervectors with ED measure in SD. Secondly, we examine the effectiveness of GMM-supervector formulation using ED measure in MD.

We use the same dataset used in our earlier work [11] to conduct experiments. The dataset is a subset of emotions of an Actress And Housewives (EAH) dataset. Dataset includes English emotion utterances portrayed by an actress and five housewives. The subset of the dataset includes four emotions: anger, happiness, neutral and sadness. The following is the data collection processes. We prepare a total of 400 emotionally neutral sentences. An actress repeated each of 400 sentences with each of four emotions. Hence, a total of  $400 \times 4 = 1600$  utterances are obtained from an actress. And, each housewife repeated each of 100 sentences with each of four emotions. Hence, a total of  $100 \times 4 \times 5 = 2000$  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 four emotions. Utterances are sampled at 16kHz and 16 bit rate.

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 and train a GMM model. Instead of training the GMM using expectation-maximization (EM) algorithm [16], we use maximum a posteriori (MAP) criterion [17] to adapt the GMM from a Universal Background Model(UBM). We train UBM via EM algorithm using 64 mixtures. We describe the details of the data used to train the UBM in the following sections. We adapt the mean and covariance only. In all GMMs, covariance is diagonal matrix. Once we have an adapted GMM model from each emotion utterance, we formulate GMM-supervectors using the techniques mentioned in sections 2 and 3. We use SVM-Torch [18] for the training and classification. We employ the target model against anti-model strategy for multi-class SVM classification.

As mentioned above, 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. In each experiment, the estimated classification accuracy is the average 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 in SD

We repeat an experiment of our earlier work [11] as experimental settings are different. In [11], we use National Institute of Standards and Technology (NIST) speaker recognition evaluation (SRE) 2001 dataset [19] to train UBM model. In this experiment, we train one UBM using all emotion utterances of training data of subset of EAH. GMM-supervector formulation is in SD space as only one UBM is employed. We change the UBM training data to compare fairly with the system using ED measure in MD.

We formulate the GMM-supervector using equation (3) which employs the Bhattacharyya distance based kernel. In addition, we also formulate the GMM-supervector using equation (2) that uses KL based kernel. We use both the first term of mean and second term of covariance in these supervector formulations. We also formulate the generalized GMM-supervectors by concatenating mean and covariance terms of an adapted UBM. 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 (%) over 6 folds cross validation*

Method	<i>GEN</i>	<i>KL</i>	<i>Bhat</i>
SD	59.28	67.19	77.91
MD	62.57	69.03	81.25

Both Bhattacharyya and KL based kernels are attributed to the approximation of dissimilarity measure between two distributions [12], [20]. The results show that GMM-supervectors that employ ED measure perform 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  $i^{th}$  subvectors in Equations (2) and (3), the latter one with Bhattacharyya kernel includes the term  $(m_i^\lambda - m_i^u)$  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. In the following section, we perform the experiments to observe the effect of employing ED measure in MD in GMM-supervector formulation.

#### 4.2. Effect of GMM-supervector with ED measure in MD

In this experiment, we formulate GMM-supervectors using the procedures mentioned in Section 3 for Bhattacharyya kernel, KL kernel and generalized GMM-supervectors. We use both mean and covariance statistics in GMM-supervector formulation. A total of 4 emotion dependent UBMs are trained for each of 3 GMM-supervector formulation techniques. Each UBM is trained using all utterances of an emotion class from training data. We perform emotion classification experiments using these 3 different GMM-supervectors and the average accuracies achieved are shown in the third row of Table 1. As we have expected, GMM-supervectors with ED measure in MD perform better than that in SD for all 3 different GMM-supervector formulation methods. We can see absolute improvement in accuracies with 3.29%, 1.84% and 3.34% for generalized GMM-supervectors, KL kernel based GMM-supervector and Bhat-

tacharyya kernel based GMM-supervector respectively. It is obvious that GMM-supervectors that employ ED measure in MD contain more information comparing with GMM-supervectors in SD. Dissimilarity information in multiple dimension helps to locate the boundaries between different emotion groups more accurately than the dissimilarity information in single dimension as illustrated in Figures 2(a) and 2(b).

One may argue that SD based emotion classification system may perform well if the number of parameters or numbers of GMM mixtures is increased. We increase numbers of GMM mixtures beyond 64 and we observe that the performance of SD system starts to deteriorate. The mean duration of utterances is only 4.8 seconds and it limits the number of parameters of SD system. However, MD system has no such limitations. We can easily add additional reference points (UBMs) as long as these points can provide additional information to the system.

#### 4.3. Analysis of confusion between different emotions

We analyzed the confusion between different emotions using the results obtained by Bhattacharyya kernel based GMM-supervector in MD. Confusion matrix is shown in Table 2. From the results, we can see that happiness emotion is misclassified with anger more often than other emotions. And, neutral samples are confused with happiness. Sadness is confused with neutral more than others. Arousal level of happiness is more similar to anger than other emotions. And, the most similar arousal level of sadness is neutral among all emotions. Hence, we can conclude that misclassification is possible between the two emotions if their arousal levels are similar.

Table 2: *Confusion matrix of emotion classification using SVM system with Bhattacharyya kernel based GMM-supervector in MD. (A: anger; H: happiness; N: neutral; S: sadness)*

		Classified Emotions			
		A	H	N	S
Actual Emotions	A	<b>91.2</b>	6.44	1.44	0.89
	H	16.89	<b>72.89</b>	5.22	5
	N	0	2.22	<b>97.78</b>	0
	S	2.22	7.67	9	<b>81.11</b>

## 5. Conclusions

We have presented an approach to formulate GMM-supervector with Emotional Dissimilarity (ED) measure in Multi-Dimensions (MD) to classify emotions. GMM-supervectors with Bhattacharyya distance based kernel can characterize Emotional Dissimilarity (ED) measure more effectively than the GMM-supervectors with KL based kernel. Experimental results show that the proposed GMM-supervector formulation in Multi-Dimensions (MD) outperforms the GMM-supervector formulations in Single Dimension (SD) for all three different GMM supervector formulation methods using Bhattacharyya distance based kernel, KL based kernel and generalized GMM supervector.

## 6. References

- [1] Yang, N., Muraleedharan, R., Kohl, J., Demirkol, I., Heinzelman, W., and Sturge-Apple, M., "Speech-based Emotion Classification Using Multiclass SVM with Hybrid Kernel and Thresholding Fusion", in proceedings of the 4th IEEE Workshop on Spoken Language Technology (SLT), Miami, Florida, December 2012.
- [2] Wu, S., Falk, T., Chan, W., "Automatic Speech Emotion Recognition Using Modulation Spectral Features". *Speech Communication* 53(5):768-785, 2011.
- [3] 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.
- [4] 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.
- [5] Pao, T.L., Chen, Y.T., Yeh, J.H., and Li, P.J., "Mandarin emotional speech recognition based on SVM and NN", 18th International Conference on Pattern Recognition, vol.1, pp.1096-1100, 2006.
- [6] Pertrushin, V. A., "System, Method And Article Of Manufacture For Detecting Emotion In Voice Signals Through Analysis Of A Plurality Of Voice Signal Parameters", United States patent No. 6,151,571, August. 31, 1999.
- [7] 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.
- [8] 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).
- [9] Krishnan, A., and Fernandez, M., "System And Method For Recognizing Emotional State From A Speech Signal", US Patent Application No: 2011/0295607, December. 1, 2011.
- [10] 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.
- [11] Nwe, T. L., Hieu, N. T., and Dilip, K. L., "Bhattacharyya Distance Based Emotional Dissimilarity Measure For Emotion Classification", in *Proc. of ICASSP*, pp. 7512-7516, Vancouver, Canada, May 2013.
- [12] 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.
- [13] 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
- [14] Mak, B., and Barnard, E., "Phone Clustering Using The Bhattacharyya Distance". *ICSLP 1996*
- [15] Schlosberg, H., "Three dimensions of emotions", *Psychological Review* 61 (1954) 8188.
- [16] 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.
- [17] 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.
- [18] Collobert, R., and Bengio, S., "SVM-Torch: Support Vector Machines for Large-Scale Regression Problems", *J. Mach. Learn. Res.*, vol. 1, pp. 143-160, 2001.
- [19] Przyboki, M., and Martin, A., "The NIST Year 2001 Speaker Recognition Evaluation Plan", <http://www.itl.nist.gov/iad/mig/tests/spk/2001/2001-spkrcevalplan-v05.9.pdf> (2001).
- [20] Hu, H., Xu, M. and Wu, W., "GMM Supervector Based SVM With Spectral Features For Speech Emotion Recognition," in *ICASSP*, Honolulu, Hawaii, USA, April 2007, pp. 413-416.