Cultural Differences and Similarities in Emotion Recognition

Vladimir Kurbalija
Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad
Trg Dositeja Obradovića 4
21000 Novi Sad, Serbia
+381 21 4852877
kurba@dmi.uns.ac.rs

Mirjana Ivanović
Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad
Trg Dositeja Obradovića 4
21000 Novi Sad, Serbia
+381 21 4852855
mira@dmi.uns.ac.rs

Miloš Radovanović
Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad
Trg Dositeja Obradovića 4
21000 Novi Sad, Serbia
+381 21 4852875
radacha@dmi.uns.ac.rs

Zoltan Geler
Department of Media Studies, Faculty of Philosophy, University of Novi Sad, Dr Zorana Đinđića 2
21000 Novi Sad, Serbia
+381 21 4853918
zoltang@ff.uns.ac.rs

Dejan Mitrović
Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad
Trg Dositeja Obradovića 4
21000 Novi Sad, Serbia
+381 21 4852875
dejan@dmi.uns.ac.rs

Weihui Dai
School of Management, Fudan University
Shanghai 200433, China
whdai@fudan.edu.cn

Weidong Zhao
School of Software, Fudan University
Shanghai 200433, China
whdai@fudan.edu.cn

ABSTRACT
The electroencephalogram (EEG) is a powerful method for investigation of different cognitive processes. Recently, EEG analysis became very popular and important, where classification of these signals stands out as one of the mostly used methodologies. Emotion recognition is one of the most challenging tasks in EEG analysis since not much is known about representation of different emotion in EEG signals. In addition, inducing of desired emotion is by itself difficult, since various individuals react differently to external stimuli (audio, video, etc.) In this paper, we will examine the similarities in emotion perception of different individuals on the basis of audio stimuli. Since some of the participants in the experiment did not understand the language of the stimuli, we will also investigate the impact of language understanding on emotion perception. This study presents some preliminary results of more complex experiments in the area of affective computing that are planned.

Categories and Subject Descriptors
• Computing methodologies--Knowledge representation and reasoning
• Applied computing--Life and medical sciences

Keywords
Artificial Intelligence, Emotion detection, Time-series mining, Classification.

1. INTRODUCTION
Emotions represent a multi-disciplinary topic traditionally studied by philosophy, psychology, sociology, neuroscience, medicine, etc. Due to recent developments in the new research direction of social and affective computing [24], emotions are now treated as an inter-disciplinary area, cross-cutting multiple subjects.

Human emotions are considered to be caused by specific situations and activities. Emotional change can form a unique subjective experience, and trigger a series of physiological responses through the nervous system. By detecting different physiological signals and external features, it is possible to compute and identify the emotional state of a subject to a certain extent.

Recently, emotion recognition and simulation have become an important research topic in man-machine communication. Emotional intelligence deals with modeling, recognition and
control of human emotions. It provides the key theory and essential basis for further studies of a new generation of intelligent information processing, intelligent services, and similar applications. One of the contemporary research areas based on emotions and corresponding theories is emotion recognition in human voice. Therefore, it is reasonable to apply speech signal processing and pattern recognition techniques, such as cepstrum analysis [4], dynamic time warping [15], and hidden Markov modeling [16], to perform different activities in this area. The ability to correctly interpret emotional signals is extremely important for the development of intelligent systems and their applications in a wide range of real-life scenarios.

Five essential domains for emotional intelligence were defined in [21]: knowing one’s emotions, managing emotions, motivating oneself, recognizing others’ emotions, and handling relationships. Therefore, emotional intelligence deals with one’s ability to perceive, understand, manage, and express emotions within oneself and in dealing with others.

Having in mind the importance of proper emotion detection, recognition and employment in different intelligent environments, it is necessary to intensify research efforts in these areas, and to perform experiments on numerous appropriate data sets.

Recently, we conducted our first real-life experiments with several Serbian and Chinese colleagues. During recordings, brain signals were measured as reactions to short vocal sentences in different emotional states, including happiness, jealousy, anger, and sadness, all pronounced by a native Mandarin speaker.

In this evaluation, we intend to investigate whether language understanding has an impact on emotion perception, and whether people of different nationalities “feel” the emotions differently. This experiment represents the starting point for further, more complex experiments in the area of affective computing. The end-goal is to obtain conclusions on cultural differences and similarities in emotion perception and recognition.

The rest of the paper is organized as follows. Related work is briefly presented in Section 2. Section 3 outlines the data collection and pre-processing methods. Section 4 presents the experimental setup. Emotion recognition based on the classification is discussed in Section 5. Finally, Section 6, draws the final conclusions of the presented work.

2. RELATED WORK

It can be said that emotion recognition from EEG signals is in its relative infancy, although significant breakthroughs have been made over the years. In an early review [5], the asymmetry in EEG signals acquired from frontal lobes is discussed, arguing that it may serve as both a moderator and a mediator of emotion- and motivation-related constructs. Despite the lack of a definitive conclusion on this subject, the discussion suggests that frontal lobes carry potentially useful information with respect to emotion recognition, which was confirmed by subsequent works [1][8][17][18].

At first, research on emotion recognition from EEG signals focused on stimuli from one source, mostly visual, but sometimes also on audio stimuli [1]. Roughly at the same time, two papers appeared that examined both types of stimuli, as well as their combination [1][18]. In [1], the influence of visual and musical stimuli on brain processing is examined, demonstrating that music can markedly enhance the emotional experience evoked by affective pictures. Research in [18] addresses five questions: (1) Is the modality of the stimulus recognizable from the recorded brain signals? (2) Are emotions recognizable from an EEG? (3) What is the influence of the modality of the stimulus on the recognition rate? (4) When using only five electrodes, what would be a good montage for emotion recognition? and (5) What features are interesting for emotion recognition? The paper provides evidence towards positive answers for questions 1 and 2, suggesting more difficulty regarding visual stimuli, and providing advice for electrode montage and feature use involving F3, F4 and Fpz channels. More recently, movie watching was used as combined stimuli for detection of positive and negative emotions in [17].

Valence and arousal dimensions of EEG signals for emotion classification are explored in detail in [8]. The dimensions do not provide satisfactory features for the task on their own. However, a combined representation using extreme values on both dimensions produced much improved results.

Features can be extracted from time series obtained from EEG signals in various ways. Successful approaches for the task of emotion recognition include the discrete wavelet transform [14], fractal dimension [12] and crossings features [19], with the support vector machines being among the most successful classifiers.

In this paper, we will explore classification accuracies of EEG signals with respect to emotion classification and to combination of emotion and native language classification. EEG signals are obtained using only audio stimuli where the same set of sentences is uttered by professional actors with different emotional overtones. The aim of this study is to explore whether different individuals “feel” the emotions in the same way and whether language understanding has an impact on emotion perception.

3. DATA PREPROCESSING: EXTRACTING FREQUENCY BANDS

Instead of analyzing EEG signals directly, the usual approach in emotion detection algorithms is to decompose them into distinct frequency bands, and then work on each band separately. Although the exact frequency ranges vary in different sources, the following set is most commonly used [13][14]: theta (4-8 Hz), alpha (8-14 Hz), beta (14-32 Hz), and gamma (32-64 Hz). Lower frequencies (delta, 1-4 Hz) are associated with babies and sleeping adults, while higher frequencies (i.e. above 64 Hz) represent noise.

Frequency bands can be extracted from raw EEG signals using band-pass audio filters [22]. These filters pass all frequencies within the given range, while discarding frequencies that fall out of the range. In practice, the filter is configured through the central frequency and the band width parameters. In case of the alpha band, for example, the central frequency would be 11 Hz, while the band width would be 3 Hz.

In order to pre-process our data and extract the required frequency bands, we used the two-pole Butterworth filter [3] as the concrete implementation of the band-pass filter. This implementation is freely available in FFmpeg, a collection of libraries and tools for multimedia processing [7]. Figure 1 shows the results of filtering a signal collected for one of the experiment subjects. The bottom-most sub-figure shows the raw EEG signal, while the middle and the top sub-figures show the theta and alpha frequency bands, respectively, that were obtained through data filtering.
4. EXPERIMENTAL SETUP

In order to perform the experiment and discover if there are differences in emotion recognition between people from different socio-cultural settings, we acquired the data using a rather standard procedure. During the data acquisition process, there were six participants: four native Mandarin speakers and two native Serbian speakers. Scalp electrodes were applied using the 10-20 system, which is one of the most widely-used international standards [9]. It typically includes 21 electrodes, with distances between adjacent electrodes representing 10% or 20% of the total front-to-back or left-to-right surfaces. The positions of electrodes are depicted graphically in Figure 2, with the alphanumeric nomenclature described in [23].

Instead of acquiring data from electrodes directly, it is common to use the so-called bipolar model, which measures the difference between a pair of electrodes [20]. Thus, to perform experiments with members of our two groups we constructed 8 channels, marked C1 to C8, as follows: C1: Fp1-T3, C2: Fp2-T4, C3: T3-O1, C4: T4-O2, C5: Fp1-C3, C6: Fp2-C4, C7: C3-O1, C8: C4-O2.

4.1 Time-series extraction

Experiments for data acquisition are performed in a relatively controlled environment, using the described electrodes. The main idea was that Chinese and Serbian participants have to listen to audio clips which express different kinds of emotions, while their brain activity during experiments was recorded using 8 channels (marked as C1 to C8) of combined electrodes fastened to their scalp. The audio clips were uniformly structured as presented in Table 1. All participants were right-handed and without serious eyesight problems.

The first six events represent several sentences performed by a Chinese actor simulating different emotions: angry, fear, happy, neutral, sad, and surprise. The same set of sentences is reproduced with appropriate intonation in all six emotions. The purpose of these events is to discover whether the way of pronunciation affects the perception of emotions regardless of speech understanding. The last four events represent different music clips (traditional, jazz, rock) whose purpose was to investigate whether the same music causes the same brain activity, or emotion, in different participants.

During the reproduction of audio clips, brain activity was recorded using 8 channels (C1 to C8). The part of recorded data for one participant is given in Table 2. Sample rate for this experiment was set to 125Hz. All 10 events were presented to each of the six participants and brain reactions were recorded. The number of samples/rows per participant varies from 21810 to 102914. The reason for this difference is that participants were not presented the same subset of sentences from a certain group/emotion.

As discussed earlier, the collected data in the raw form is not suitable for further meaningful analysis. The EEG signals from each channel are firstly decomposed to their alpha, beta, gamma and theta frequency band using the two-pole Butterworth filter. To obtain a meaningful form of time series appropriate for further analysis, each signal from each channel and each frequency band is split into 10 time series on the basis of the performed event.

Table 1. The specification of audio clips.

<table>
<thead>
<tr>
<th>Event Number</th>
<th>Stimulating Signal</th>
<th>Language</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Speech (angry)</td>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Speech (fear)</td>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Speech (happy)</td>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Speech (neutral)</td>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Speech (sad)</td>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Speech (surprise)</td>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Music (Chinese classic song)</td>
<td>Mandarin</td>
<td>The joy of love</td>
</tr>
<tr>
<td>8</td>
<td>Music (jazz)</td>
<td>English</td>
<td>Fly me to the moon</td>
</tr>
<tr>
<td>9</td>
<td>Music (Chinese traditional instruments)</td>
<td>Mandarin</td>
<td>Butterfly Lovers</td>
</tr>
<tr>
<td>10</td>
<td>Music (rock and roll)</td>
<td>English</td>
<td>Breed</td>
</tr>
</tbody>
</table>

CLASSIFICATION ACCURACIES are performed using the FAP system [10].

were z-normalized before the experiments. All the experiments were chosen as one of the most prominent in time-series mining.

Another reason for the choice of DTW is its natural fit to the time series data [6]. As the distance measure for this distance classifier was used since it was shown that it gives among the best results (compared to many not only distance-based classifiers) with time-series data [6][11]. For the evaluation technique we used 10 runs of 10-fold stratified cross validation (SCV10x10). The simple 1NN classifier was used in the first set of experiments, we measure classification error rate on the prepared time-series dataset, where the labels of time series consist only of events (6 emotions and 4 kinds of music). The original dataset is divided into 32 sub-datasets, where each sub-dataset contains time series with the same channel and frequency band. Each sub-dataset contains 60 time series: 6 participants with 10 events. In each sub-dataset, SCV10x10 is performed with a 10-class classifier. The purpose of this experiment was to eventually determine which combination of channel and frequency band gives the best results in emotion recognition. The good results would be indicated by relatively small classification error rate. The results for all 32 sub-datasets are given in Table 3.

In the first phase of the analysis, we have incorporated into the class labels the information about the nationality of the participants. This could indicate that different people experience the same emotion in similar way, independently of their nationality or language understanding.

Finally, we obtained a labeled dataset which consists of 1920 time series. The label of each time series is constructed in two different ways:

1. Event number (emotion) from Table 1.
2. A combination of the event number and the nationality of the participant.

5. EMOTION RECOGNITION BASED ON CLASSIFICATION ACCURACIES

With the time series prepared in an adequate way we performed two sets of experiments. The experimental setup was the same for both sets, while only the used labels of time series were different.

In the first set of the investigations, the time series were labeled relying solely on the emotion (the event number) regardless of the participants. The aim of these experiments was to examine whether there are similarities or common patterns between the time series of different participants for the same emotion. This could indicate that different people experience the same emotion in similar way, independently of their nationality or language understanding.

In the second phase of the analysis, we have incorporated into the class labels the information about the nationality of the participants, too. The motivation for this extension was in the assumption that there may be significant differences in the responses of the participants of different nationalities which lead to considerably dissimilar time series.

The common methodology for estimating the similarities between time series is to measure classification accuracy of labeled time series [6][11]. For the evaluation technique we used 10 runs of 10-fold stratified cross validation (SCV10x10). The simple 1NN classifier was used since it was shown that it gives among the best results (compared to many not only distance-based classifiers) with time-series data [6]. As the distance measure for this distance based classifier, the dynamic time warping (DTW) [2] measure was chosen as one of the most prominent in time-series mining. Another reason for the choice of DTW is its natural fit to the scenario where time series are of different length. The time series were z-normalized before the experiments. All the experiments are performed using the FAP system [10].

<table>
<thead>
<tr>
<th>Event Number</th>
<th>Min. Sec.</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.241</td>
<td>74.453</td>
<td>34.601</td>
<td>-9.129</td>
<td>41.927</td>
<td>-11.496</td>
<td>63.229</td>
<td>-6.424</td>
<td>33.474</td>
</tr>
<tr>
<td>1</td>
<td>1.241</td>
<td>74.461</td>
<td>19.048</td>
<td>-9.129</td>
<td>41.927</td>
<td>-11.496</td>
<td>63.229</td>
<td>-6.424</td>
<td>33.474</td>
</tr>
<tr>
<td>1</td>
<td>1.242</td>
<td>74.492</td>
<td>14.765</td>
<td>3.719</td>
<td>30.769</td>
<td>10.144</td>
<td>32.122</td>
<td>1.014</td>
<td>23.331</td>
</tr>
<tr>
<td>1</td>
<td>1.242</td>
<td>74.5</td>
<td>-902</td>
<td>-37.87</td>
<td>20.626</td>
<td>-17.244</td>
<td>13.187</td>
<td>-25.359</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. The structure of raw data for one participant.

Table 3. The results of classification with respect to emotion.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Wave</th>
<th>Err. Rate</th>
<th>Channel</th>
<th>Wave</th>
<th>Err. Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>alpha</td>
<td>0.93</td>
<td>C5</td>
<td>alpha</td>
<td>0.918333</td>
</tr>
<tr>
<td>C1</td>
<td>beta</td>
<td>0.963333</td>
<td>C5</td>
<td>beta</td>
<td>0.945</td>
</tr>
<tr>
<td>C1</td>
<td>gamma</td>
<td>0.925</td>
<td>C5</td>
<td>gamma</td>
<td>0.926667</td>
</tr>
<tr>
<td>C1</td>
<td>theta</td>
<td>0.938333</td>
<td>C5</td>
<td>theta</td>
<td>0.878333</td>
</tr>
<tr>
<td>C2</td>
<td>alpha</td>
<td>0.905</td>
<td>C6</td>
<td>alpha</td>
<td>0.938333</td>
</tr>
<tr>
<td>C2</td>
<td>beta</td>
<td>0.958333</td>
<td>C6</td>
<td>beta</td>
<td>0.896667</td>
</tr>
<tr>
<td>C2</td>
<td>gamma</td>
<td>0.928333</td>
<td>C6</td>
<td>gamma</td>
<td>0.926667</td>
</tr>
<tr>
<td>C2</td>
<td>theta</td>
<td>0.911667</td>
<td>C6</td>
<td>theta</td>
<td>0.911667</td>
</tr>
<tr>
<td>C3</td>
<td>alpha</td>
<td>0.876667</td>
<td>C7</td>
<td>alpha</td>
<td>0.908333</td>
</tr>
<tr>
<td>C3</td>
<td>beta</td>
<td>0.975</td>
<td>C7</td>
<td>beta</td>
<td>0.956667</td>
</tr>
<tr>
<td>C3</td>
<td>gamma</td>
<td>0.893333</td>
<td>C7</td>
<td>gamma</td>
<td>0.928333</td>
</tr>
<tr>
<td>C3</td>
<td>theta</td>
<td>0.871667</td>
<td>C7</td>
<td>theta</td>
<td>0.921667</td>
</tr>
<tr>
<td>C4</td>
<td>alpha</td>
<td>0.923333</td>
<td>C8</td>
<td>alpha</td>
<td>0.88</td>
</tr>
<tr>
<td>C4</td>
<td>beta</td>
<td>0.878333</td>
<td>C8</td>
<td>beta</td>
<td>0.913333</td>
</tr>
<tr>
<td>C4</td>
<td>gamma</td>
<td>0.923333</td>
<td>C8</td>
<td>gamma</td>
<td>0.94</td>
</tr>
<tr>
<td>C4</td>
<td>theta</td>
<td>0.908333</td>
<td>C8</td>
<td>theta</td>
<td>0.893333</td>
</tr>
</tbody>
</table>
The error rate of a 10-class random classifier with equal distribution of classes is 0.9. The presented results show that all the combinations are around the level of a random classifier, and no one combination of channel and frequency band stands out. This means that the most similar time series from the training set (according to DTW) most often do not belong to the same class. In other words, none of these combinations can be used to effectively distinguish between recordings of the participants’ brain activity induced by listening to audio clips which express different kinds of emotions.

In the second set of experiments the division into 32 sub-datasets is the same as in the first setup. The only difference is that classification is performed according to events and nationality of participant. This represents 20-class classification (2 nationalities: Chinese and Serbian with 10 events). In this setup, the labels are not equally distributed: each sub-dataset contains recordings of four Chinese and two Serbian colleagues.

By this evaluation we intended to investigate whether language understanding has an impact on emotion perception and whether different nationalities “feel” the emotions differently. The results are given in Table 4.

Table 4. The results of classification with respect to emotion and nationality.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Wave</th>
<th>Err. Rate</th>
<th>Channel</th>
<th>Wave</th>
<th>Err. Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>alpha</td>
<td>0.931667</td>
<td>C5</td>
<td>alpha</td>
<td>0.941667</td>
</tr>
<tr>
<td>C1</td>
<td>beta</td>
<td>0.976667</td>
<td>C5</td>
<td>beta</td>
<td>0.94</td>
</tr>
<tr>
<td>C1</td>
<td>theta</td>
<td>0.976667</td>
<td>C5</td>
<td>theta</td>
<td>0.896667</td>
</tr>
<tr>
<td>C1</td>
<td>gamma</td>
<td>0.993333</td>
<td>C5</td>
<td>gamma</td>
<td>0.998</td>
</tr>
<tr>
<td>C2</td>
<td>alpha</td>
<td>0.946667</td>
<td>C6</td>
<td>alpha</td>
<td>0.941667</td>
</tr>
<tr>
<td>C2</td>
<td>beta</td>
<td>0.961667</td>
<td>C6</td>
<td>beta</td>
<td>0.896667</td>
</tr>
<tr>
<td>C2</td>
<td>theta</td>
<td>0.93</td>
<td>C6</td>
<td>theta</td>
<td>0.946667</td>
</tr>
<tr>
<td>C2</td>
<td>gamma</td>
<td>0.961667</td>
<td>C6</td>
<td>gamma</td>
<td>0.978333</td>
</tr>
<tr>
<td>C3</td>
<td>beta</td>
<td>0.971667</td>
<td>C7</td>
<td>beta</td>
<td>0.961667</td>
</tr>
<tr>
<td>C3</td>
<td>gamma</td>
<td>0.961667</td>
<td>C7</td>
<td>gamma</td>
<td>0.995</td>
</tr>
<tr>
<td>C3</td>
<td>theta</td>
<td>0.895</td>
<td>C7</td>
<td>theta</td>
<td>0.923333</td>
</tr>
<tr>
<td>C4</td>
<td>alpha</td>
<td>0.963333</td>
<td>C8</td>
<td>alpha</td>
<td>0.931667</td>
</tr>
<tr>
<td>C4</td>
<td>beta</td>
<td>0.911667</td>
<td>C8</td>
<td>beta</td>
<td>0.928333</td>
</tr>
<tr>
<td>C4</td>
<td>gamma</td>
<td>0.981667</td>
<td>C8</td>
<td>gamma</td>
<td>0.998333</td>
</tr>
<tr>
<td>C4</td>
<td>theta</td>
<td>0.926667</td>
<td>C8</td>
<td>theta</td>
<td>0.926667</td>
</tr>
</tbody>
</table>

The results are again around the level of a random classifier (for a 20-class random classifier with equal distribution of classes the error rate is 0.95) with no specific emphasized combination. This means that even in this case no combination can be successfully used to differentiate time series corresponding to distinct classes. The obtained classification accuracies may indicate that the brain signals recorded under the influence of listening to audio tracks of different emotions are not considerably different (according to DTW).

Some potentially useful findings from the perspective of emotion detection could be:

- Channel C3 in most of the cases gives better results,
- Frequency band gamma generally gives worst results.

However, these findings are not formally grounded, and more investigation is needed to prove them.

Within the experiments described in this paper we have investigated the possibility of applying some techniques of time-series data mining in the field of emotion recognition in human voice. Participants of different nationalities have listened to audio clips which express different emotions. After decomposing the recorded brain signals into alpha, beta, gamma and theta frequency bands, we analyzed them in an attempt to discover similarities between recordings which correspond to the same emotions. The first setup of the experiments did not differentiate the recordings on the basis of the participants’ nationality. The second approach has taken into account this information, too. However, neither of these two approaches was able to identify a combination of channel and frequency band that would help in distinguishing between brain signals recorded while listening to audio clips with different emotions.

6. CONCLUSION

Although the results presented in this paper are primarily rather negative, some conclusions can still be drawn from them. Firstly, the instances comprising the data sets on which experiments were run all originate from different individuals. Therefore, it appears that signals corresponding to the same emotion produced by different people are very different (at least from the viewpoint of DTW distance). Secondly, in the literature suggests that channels associated to the frontal lobe carry most useful information with respect to the emotion recognition task. In our measurements the frontal lobe was only partially accounted for, which may be another reason for the exhibited high error rates. In future work we will construct data sets containing multiple instances of signals that originate from one individual, with the expected effect of achieving much better classification accuracy.

In future work we will also investigate discrimination of nationality based on the same ECG signals. Preliminary results are positive, indicating that the use of some channels and bands which were not adequate for emotion recognition could carry useful information for classification based on nationality. This suggests that the language barrier may have a useful role with respect to this task.

7. ACKNOWLEDGMENTS

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, through project no. OI174023: “Intelligent techniques and their integration into wide-spectrum decision support,” and by bilateral project no. 2-9/2013: “Emotional Intelligence and Applications Based on Multi-Agent Systems.”

8. REFERENCES


