ISSN 2229-5518

Characterization of Human Affective States Using Multichannel Multiscale Entropy (MMSE) Analysis Method

Kawser Ahammed

Abstract—This work is based on the recently introduced multivariate multiscale entropy (MMSE) analysis method. In this article, MMSE analysis method is applied on a multimodal dataset provided by Sander Koelstra et al. for analyzing emotions from physiological signals in order to characterize human affective states. The multimodal dataset contains electroencephalogram (EEG) and peripheral physiological signals recorded from 32 participants while the participants were watching 40 selected music videos of one minute duration. Each participant rated their emotional response to 40 music videos along the scales of valence, arousal, dominance and liking. The MMSE analysis curve obtained using this multimodal dataset shows differences in terms of complexity among different affective states, which can be used for emotion detection and classification for machine vision applications.

Keywords—Emotion classification, EEG, Physiological signals, MMSE, multimodal dataset, machine vision applications, Affective states computing

1 INTRODUCTION

Psycho-physiological process triggered by conscious and/or unconscious perception of an object or situation is called emotion. In other words, an emotion is usually caused by a person consciously or unconsciously evaluating an event as relevant to a concern that is important and is usually experienced as a distinctive type of mental state, sometimes accompanied or followed by bodily changes, expressions, actions; the emotion is felt as positive when a concern is advanced and negative when a concern is impeded [1]. Emotion is often related to mood, temperament, personality and disposition, and motivation. Emotions contribute a significant role in human communication. Emotions can be expressed either verbally through emotional vocabulary, or by expressing non-verbal cues such as intonation of voice, facial expressions and gestures. Most of the contemporary human-computer interaction (HCI) systems are unable to identify human affective states. So, the goal of characterization of human affective states is to fill this gap by detecting human affective states occurring during human-computer interaction in terms of complexity [2].

Various discrete categorizations of emotions have been proposed, such as the six basic emotions proposed by Ekman and Friesen [3] and the tree structure of emotions proposed by Parrot [4]. Among six basic emotions, arousal can range from inactive (e.g. uninterested, bored) to active (e.g. alert, excited), whereas valence ranges from unpleasant (e.g. sad, stressed) to pleasant (e.g. happy, elated). Dominance ranges from a helpless and weak feeling (without control) to an empowered feeling (in control of everything). Emotion assessment is often carried out through analysis of user's emotional expressions and/or physiological signals. Emotional expressions refer to any observable verbal and non-verbal behavior that communicates emotion. So far, most of the studies on emotion assessment have focused on the analysis of facial expressions and speech to determine a person's emotional state. Emotional information that can be used for emotion assessment exists in physiological signals.

The creation of novel databases containing emotional expressions in different modalities has been motivated by recent developments in emotion recognition. The contents of these databases are speech, visual or audiovisual data [5], [6], [7], [8], [9]. Facial expressions and/or body gestures belong to the visual modality. Posed or genuine emotional speech in different languages is included in the audio modality.

Multichannel multiscale entropy (MMSE) analysis method proposed by Mosabber *et al.* [10] measures the complexity of multichannel data. The possible applications of this method are analysis of brain consciousness [11], complexity analysis of multichannel data [12], dynamical complexity of human responses [13], and human centred complexity analysis [14]. This method never analysed human emotions.

In this work, a multimodal database containing emotional data has been used to analyse human emotions. To make that database, S. Koelstra *et al.* [2] used 40 music video clips as the visual stimuli to elicit different emotions. Using these emotional data, multichannel sample entropy is calculated in terms of scale factor in order to determine higher complexity emotional state. TABLE 1 gives an overview of the database contents.

• *Kawser Ahammed, MS Scholar, Dept. of Electrical & Electronic Engineering (Former: Applied Physics, Electronics and Communication Engineering), University of Dhaka, Bangladesh, PH-8801731955348, E-mail: kawser. apece@gmail.com*

TABLE 1

Database content summary, adapted from Ref. [2]

Online subjective annotation	
Number of videos	120
Video duration	1 minute affective highlight
Selection method	60 via last.fm affective tags, 60 manually selected
No. of ratings per video	14-16
Ratings scales	Arousal, Valence, Dominance, liking
Ratings values	Discrete scale of 1-9
Physiological experiment	
Number of participants	32
Number of videos	40
Selection method	Subset of online annotated videos with clearest responses
Ratings scales	Arousal, Valence, Dominance, Liking, Familiarity
Ratings values	Familiarity: discrete scale of 1-5, others: continuous scale of 1-9
Recorded signals	32 channel 512 Hz EEG, Peripheral physiological signals, face video

2 DEAP: A DATASET FOR EMOTION ANALYSIS

The DEAP dataset consists of two parts [15]:

(i). The ratings from an online self-assessment where 120 one-minute extracts of music videos were each rated by 14-16 volunteers based on arousal, valence and dominance.

(ii). The participant ratings, physiological recordings and face video of an experiment where 32 volunteers watched a subset of 40 of the above music videos EEG and physiological signals were recorded and each participant also rated.

3 DATA PREPROCESSED MATLAB. ZIP

Recordings of EEG and peripheral physiological signals were downsampled and were kept in zip files. These files contain a downsampled (to 128Hz), preprocessed and segmented version of the data in Matlab. Each zip file contains 32 .mat (matlab) files, one per participant. Each participant file contains two arrays [15]:

Array name	Array shape	Array contents
Data	40×40 ×8064	video/trial×channel ×data
Labels	40×4	video/trial ×labels (valence, arousal, dominance, liking)

TABLE 2

Channel layout, adapted from Ref. [15]

Channel no.	Channel content
33	hEOG (horizontal EOG, hEOG1 - hEOG2)
34	vEOG (vertical EOG, vEOG1 - vEOG2)
35	zEMG (Zygomaticus Major EMG, zEMG1 - zEMG2)
36	tEMG (Trapezius EMG, tEMG1 - tEMG2)
37	GSR (Ohm)
38	Respiration belt
39	Plethysmograph
40	Temperature

TABLE 2 shows the channel layout of 8 channels among 40 channels. In this article, 33-40 channels have been utilized for emotion analysis except 37 and 38 channels. Emotion analysis has been performed using these channels in terms of complexity. If an emotion (e.g. valence) has higher multivariate sample entropy than the other (e.g. arousal), valence emotion will have higher complexity than that of arousal. This multivariate sample entropy theory proposed by Mosabber *et al.* [10] has been used in this article to measure the complexity of emotion.

ISSN 2229-5518

TABLE 3

Labels Array for a participant among 32 participants, adapted from Ref. [16]

40(video/trial)×4 (labels)

Video/trial	Labels			
	Valence	Arousal	Dominance	Liking
1	9.0	5.03	7.13	6.62
2	8.01	7.1	8.04	8.08
3	9.0	9.0	8.96	8.05
4	6.05	1.0	5.04	7.03
5	5.04	3.0	3.65	5.04
6	5.0	4.94	5.04	1.01
7	4.96	1.99	4.08	1.0
8	9.0	9.0	9.0	9.0
9	9.0	7.0	9.0	7.18
10	4.99	1.0	9.0	1.0
11	7.08	1.0	7.01	7.0
12	8.01	6.06	7.12	9.0
13	9.0	4.99	9.0	9.0
14	8.94	9.0	8.97	9.0
15	9.0	8.01	8.06	9.0
16	6.0	5.05	8.05	6.14
17	9.0	3.0	6.03	7.41
18	9.0	6.15	9.0	8.08
19	9.0	8.32	8.06	6.99
20	8.97	9.0	8.03	9.0
21	1.0	1.0	5.04	1.0
22	5.0	1.0	4.97	5.0
23	1.0	2.97	1.0	7.82
24	7.08	9.0	6.06	9.0
25	9.0	9.0	8.1	9.0
26	9.0	8.04	9.0	9.0
27	4.96	5.01	3.01	5.05
28	4.06	4.97	7.14	7.04
29	2.99	7.04	7.03	9.0
30	2.01	7.05	1.0	7.13
31	4.06	7.05	1.0	9.0
32	1.0	9.0	2.01	1.0
33	5.0	8.03	3.01	1.0
34	4.87	1.0	6.99	1.0
35	1.99	2.0	1.0	1.0
36	5.04	1.0	7.09	1.0
37	1.0	8.06	1.0	1.0
38	1.0	9.0	1.0	1.0
39	1.0	1.0	2.97	1.0
40	6.05	6.67	6.94	6.53

4 MMSE Method

The multichannel multiscale entropy (MMSE) method is performed through the following two steps:

(i). Define temporal scales of increasing length by coarse graining the multichannel time series $\{u_{l,i}\}_{i=1}^{N}$, l=1, 2, . . . ,c, where c denotes the number of channels and N the number of samples in each channel. Then, for a scale factor, ε, the elements of the multichannel coarse-grained time series are calculated as

$$x_{l,j}^{\varepsilon} = \frac{1}{\varepsilon}\sum_{i=(j-1)\varepsilon+1}^{j\varepsilon} u_{l,i} \tag{1}$$

Where $1 \leq j \leq \frac{N}{\varepsilon}$ and $l = 1, \dots, c$.

(ii). Calculate the multichannel sample entropy, MSampEn, for each coarse-grained multichannel, $x_{l,j}^{\varepsilon}$, and plot MSampEn as a function of the scale factor ε. Multichannel sample entropy is therefore a prerequisite for performing multiscale entropy (MSE) analysis simultaneously over a number of data channels.

5 Multichannel Sample Entropy Calculation

For a c-channel time series $\{u_{l,i}\}_{i=1}^{N}$, l=1, 2, . . . , c, Multichannel Sample Entropy is performed through the following steps:

i. Form $(N-n)$ composite delay vectors $U_m(i) \epsilon \mathbb{R}^m$, where $i = 1, 2, \dots, N-n$ and $n = max\{M\} \times \max\{\tau\}$

ii. Define the distance between any two composite delay vectors $U_m(i)$ and $U_m(j)$ as the maximum norm, $d[U_m(i)\ and\ U_m(j)]\ max_{k=1,\dots,m}\ \{|u(i+k-1) - u(j+k-1)|\}$

iii. For a given composite delay vector $U_m(i)$ and a threshold r, count the number of instances C_i where $d[U_m(i), U_m(j)] \leq r, j \neq i$, then calculate the frequency of occurrence, $A_i^m(r) = \frac{1}{N-n-1} C_i$, and define a global quantity

$$A^m(r) = \frac{1}{N-n}\sum_{i=1}^{N-n} A_i^m(r) \tag{2}$$

iv. Extend the dimensionality of the multichannel delay vector from m to $(m+1)$. This can be performed in c different ways, as from a space defined by the embedding vector $M[m_1, m_2, \dots, m_l, \dots, m_c]$ the system can evolve to any space for which the embedding vector is $[m_1, m_2, \dots, m_l + 1, \dots, m_c] (l = 1,2, \dots, c)$. Thus, a total of $c \times (N-n)$ vectors $U_{m+1}(i)$ in $\mathbb{R}^{m+1}$ are obtained, where $U_{m+1}(i)$ denotes any embedded vector upon increasing the embedding dimension from m_l to $(m_l + 1)$ for a specific variable l. In the process, the embedding dimension of the other data channels is kept unchanged, so that the overall embedding dimension of the system undergoes the change from m to $(m+1)$.

v. For a given $U_{m+1}(i)$, calculate the number of vectors Q_i, such that $d[U_{m+1}(i), U_{m+1}(j)] \leq r$, where $j \neq i$, then calculate the frequency of occurrence, $A_i^m(r) = \frac{1}{c(N-n)-1} Q_i$, and define the global quantity

$$A^{m+1}(r) = \frac{1}{c(N-n)} \sum_{i=1}^{c(N-n)} A_i^{m+1}(r) \tag{3}$$

vi. This way, $A^m(r)$ represents the probability that any two composite delay vectors are similar in dimension m, whereas $A^{m+1}(r)$ is the probability that any two composite delay vectors will be similar in dimension $(m+1)$.

vii. Finally, for a tolerance level r, MSampEn is calculated as the negative of a natural logarithm of the conditional probability that two composite delay vectors close to each other in an m dimensional space will al so be close to each other when the dimensionality is increased by one, and is given by

$$MS_{En}(M, \tau, r, N) = -\ln\left[\frac{A^{m+1}(r)}{A^m(r)}\right] \tag{4}$$

Where the symbol , MS_{En}, denotes Multichannel Sample Entropy

6 Result and Analysis

6.1 Complexity analysis of physiological signals for emotion analysis

For complexity analysis the highest values (e.g. 9, 8.05, 8.06, 8.08, 8.03) of each emotion label are used among the four emotion labels. The highest values of each emotion label are given in TABLE 4.

TABLE 4
Highest values for four emotion labels, adapted from Ref. [16]
13(video/trial, among 40 videos) ×4 (labels)

Video/trial	Labels			
	Valence	Arousal	Dominance	Liking
1	9.0			
2				8.08
10			9.0	
12				9.0
16			8.05	
17	9.0			
19	9.0			
23				7.82
28			7.14	
29				9.0
31				9.0
32		9.0		
33		8.03		
37		8.06		
38		9.0		

From the TABLE 4, it is obsrerved that for the first trial arousal, dominance and liking values are absent. This is because of

ISSN 2229-5518

choosing the highest value for any emotion label among valence, arousal, dominance and liking emotion labels. From TABLE 3, it can be shown that for the first trial valence= 9.0, arousal= 5.03, dominance= 7.13 and liking= 6.62. Therefore, the highest value is obtained by the valence emotion label. Similarly, other highest values of emotion are achieved.

6.2 Complexity analysis of multichannel data for emotion analysis

In this section, hEOG-vEOG, zEMG-tEMG and Plethysmograph-Temperature multichannel contents are analyzed in terms of complexity in order to test affective states of human. The values of the parameters used to calculate multichannel sample entropy were m_1 = 3, m_2=2, τ_1=1, τ_2=2 and r is set at a certain percentage (usually 7%) of the original channel content standard deviation (SD).

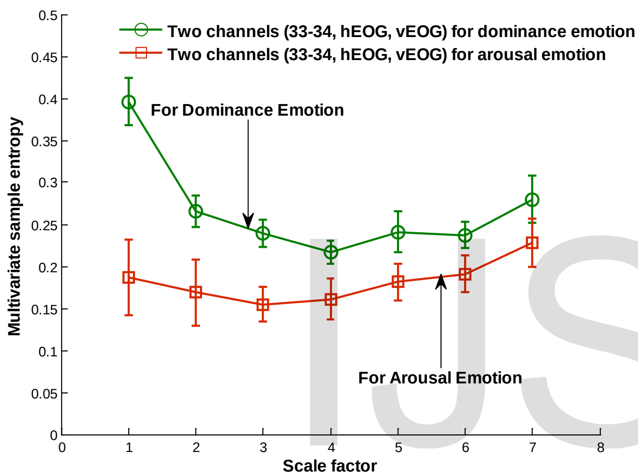
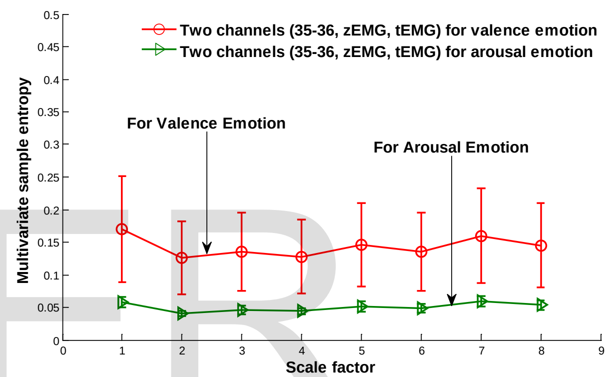


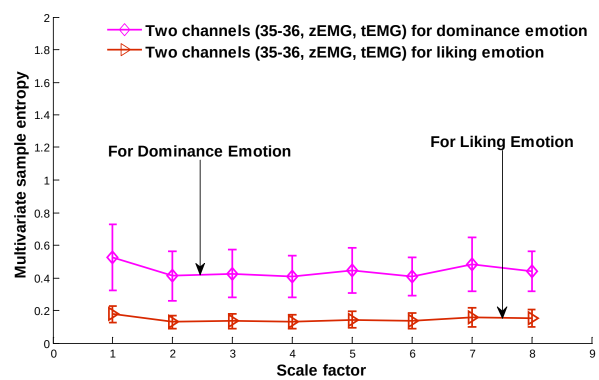
Fig. 1. MMSE (Multichannel Multiscale Entropy) analysis for two channels containing hEOG (horizontal EOG, hEOG1 - hEOG2) and vEOG (vertical EOG, vEOG1 - vEOG2) time series (mean zero, variance one), each with 7681 data points. Symbols represent an average of 4 independent trials for arousal and an average of 3 independent trials for dominance. Error bars represent the standard deviation (SD). Lines represent numerical evaluation of analytic Multichannel Sample Entropy Calculation.

Fig. 1. shows MMSE (Multichannel Multiscale Entropy) analysis for two channels containing hEOG (horizontal EOG, hEOG1 - hEOG2) and vEOG (vertical EOG, vEOG1 - vEOG2) time series (mean zero, variance one), each with 7681 data points. From fig. 1, it is observed that the values of sample entropy for the dominance emotion are higher than the values of sample entropy for the arousal emotion for all scales. Due to the higher entropy values for all scales for dominance rather than arousal, this result is consistent with the fact that, unlike arousal, dominance contains correlations across multiple time scales and is, therefore, more complex than arousal. This fact also implies that human has an empowered feeling (in control of everything) for dominance and an inactive feeling (e.g. uninterested, bored) for arousal.

Fig. 2. shows MMSE (Multichannel Multiscale Entropy) analysis for two channels containing zEMG (Zygomaticus Major EMG, zEMG1 - zEMG2) and tEMG (Trapezius EMG, tEMG1 - tEMG2) time series (mean zero, variance one), each with 7681 data points. In fig. 2. (a), it has been shown that the values of sample entropy for the valence emotion are higher than the values of sample entropy for the arousal emotion for all scales. Since the entropy values are higher for all scale factors, this result is consistent with the fact that, unlike arousal, valence emotion contains correlations across multiple time scales and is, therefore, more complex than arousal . This fact also implies that human has a pleasant feeling (e.g. happy, elated) for valence and an inactive feeling (e.g. uninterested, bored) for arousal. Similarly, from Fig. 2. (b) anyone can assume that dominance emotion is more complex than liking emotion for the same channels described in fig. 2. (a). Therefore, human has an empowered feeling (in control of everything) for dominance and a weak liking feeling (e.g. how large do you like anything) for liking emotion.



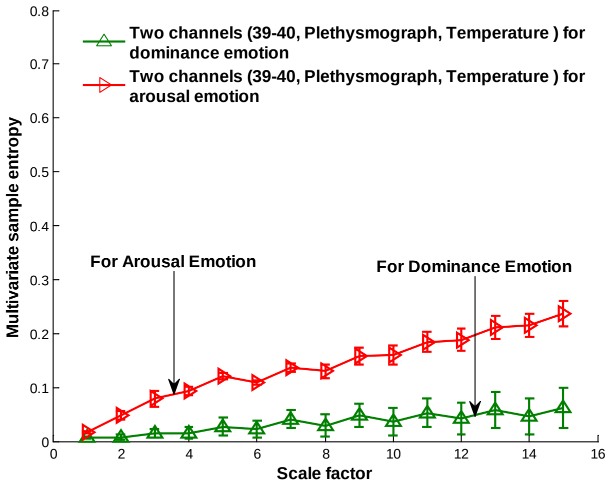
(a) For valence and arousal emotion



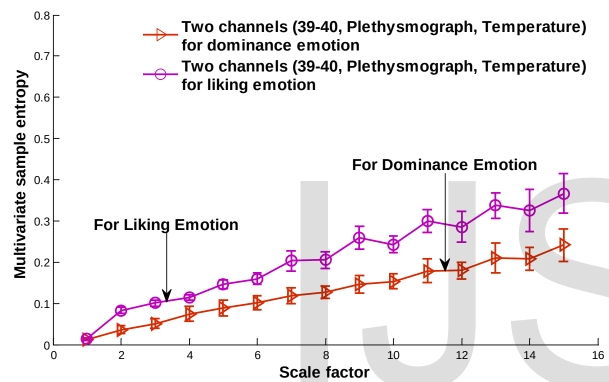
(b) For dominance and liking emotion

Fig. 2. MMSE (Multichannel Multiscale Entropy) analysis for two channels containing zEMG (Zygomaticus Major EMG, zEMG1 - zEMG2)) and tEMG (Trapezius EMG, tEMG1 - tEMG2) time series (mean zero, variance one), each with 7681 data points. Symbols represent an average of 4 independent trials for arousal, 3 independent trials for valence, 3 independent trials for dominance and 5 independent trials for liking. Error bars represent the standard deviation (SD). Lines represent numerical evaluation of analytic Multichannel Sample Entropy Calculation.

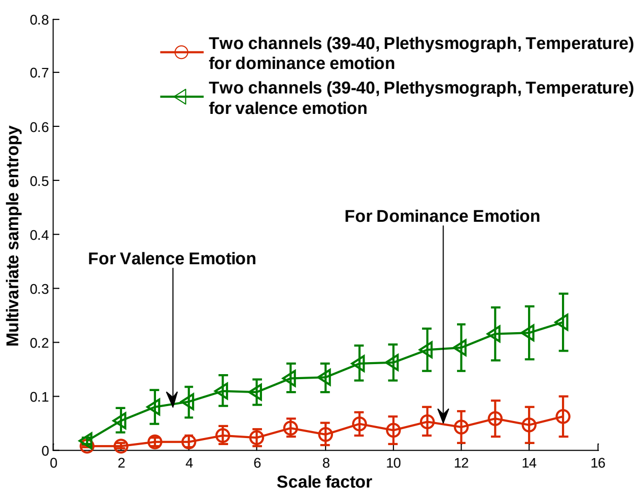
ISSN 2229-5518



(c) For arousal and dominance emotion



(d) For dominance and liking emotion



(e) For valence and dominance emotion

Fig. 3. MMSE (Multichannel Multiscale Entropy) analysis for two channels containing Plethysmograph and Temperature time series (mean zero, variance one), each with 7681 data points. Symbols represent an average of 4 independent trials for arousal, 3 independent trials for valence, 3 independent trials for dominance and 5 independent trials for liking. Error bars represent the standard deviation (SD). Lines represent numerical evaluation of analytic Multichannel Sample Entropy calculation.

Fig. 3. Shows MMSE (Multichannel Multiscale Entropy) analysis for two channels containing Plethysmograph and and Temperature time series (mean zero, variance one), each with 7681 data points. From fig. 3. (c), one can say that the values of sample entropy for the dominance and arousal emotion increase with the increase of scale factors except for scale factor 1. Both of emotions do not overlap with each other except scale factor 1. This indicates that both emotions can be identified by MMSE algorithm. Since the values of sample entropy for the arousal emotion are higher than the dominance except for scale factor 1, it can be said that, unlike dominance, arousal emotion contains correlations across multiple time scales and is, therefore, more complex than dominance . This fact also implies that human has an active (e.g. alert, excited) feeling for arousal and helpless, weak (without control) feeling for dominance.

In fig. 3. (d), it is shown that the values of sample entropy for dominance and liking emotion do not overlap with each other except scale factors 1 and 4. Because of no overlapping behavior for the majority of the scale factors, both emotions can be identified by MMSE algorithm. In this case, Liking emotion is more complex than the dominance. This fact also implies that human has a strong liking feeling for liking emotion and helpless, weak (without control) feeling for dominance.

In fig. 3. (e), one can see that for scale factor ≥ 2, the values of sample entropy for the dominance and valence emotions increase. Both of them do not overlap with each other except scale factor 1. This indicates that both emotions can be identified by MMSE algorithm. Since the values of sample entropy for the valence emotion are higher than the dominance, it can be said that, unlike dominance, valence emotion contains correlations across multiple time scales and is, therefore, more complex than dominance. This fact also implies that human has a pleasant feeling (e.g. happy, elated) for valence and helpless, weak (without control) feeling for dominance.

7 Conclusion

This work applies the recent introduced multichannel multiscale entropy analysis method on a multimodal dataset provided by Sander Koelstra et al. for the analysis of spontaneous emotions using physiological signals. Most of the contemporary human-computer interface action (HCI) systems are unable to identify human emotional states. So the goal of this work is to resolve the lack of emotional intelligence suffered from the human-computer interaction systems. In this work, multichannel sample entropies are calculated for different emotions such as valence, arousal, and dominance using physiological signals and then are compared to measure which emotion has the higher complexity than the other in order to detect human affective states.

ACKNOWLEDGEMENT

I would like to express my gratitude to my thesis supervisor Dr. Mosabber Uddin Ahmed from Electrical & Electronic Engineering – University of Dhaka, Bangladesh. I am heartily thankful to him for his constant support and guidelines.

REFERENCES

[1] Oatley, K. & Jenkins, J.M. Malden. "Understanding Emotions", Blackwell Publishers, Oxford, UK, 1996

[2] S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, "DEAP: A Database for Emotion Analysis using Physiological Signals", IEEE Transaction on Affective Computing, vol. 3, no. 1, pp. 1-15, Jan/March 2012, DOI: 10.1109/T-AFFC.2011.15 3.01

[3] P. Ekman, W. V. Friesen, M. OSullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, and P. E. Ricci-Bitti, "Universals and cultural differences in the judgments of facial expressions of emotion", Journal of Personality and Social Psychology, vol. 53, no. 4, pp. 712-717, Oct. 1987

[4] W. G. Parrott, "Emotions in Social Psychology: Essential Readings", Philadelphia: Psychology Press, 2001

[5] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis", in Proc. Int. Conf. Multimedia and Expo, Amsterdam, The Netherlands, 2005, pp. 317-321

[6] E. Douglas-Cowie, R. Cowie, and M. Schr oder, "A new emotion database: Considerations, sources and scope", in Proc. ISCA Workshop on Speech and Emotion, 2000, pp. 39-44

[7] H. Gunes and M. Piccardi, "A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior", in Proc. Int. Conf. Pattern Recognition, vol. 1, 2006, pp. 1148-1153

[8] G. Fanelli, J. Gall, H. Romsdorfer, T. Weise, and L. Van Gool, "A 3-D audio-visual corpus of affective communication, IEEE Trans. Multimedia, vol. 12, no. 6, pp. 591-598, Oct. 2010

[9] M. Grimm, K. Kroschel, and S. Narayanan, "The vera am mittag german audio-visual emotional speech database", in Proc. Int. Conf. Multimedia and Expo, 2008, pp. 865-868

[10] Mosabber Uddin Ahmed, Danilo P. Mandic,. "Phys. Rev. E 84'', 061918, 2011

[11] M. U. Ahmed, L. Li, J. Cao, and D. P. Mandic, "Multivariate Multiscale Entropy for Brain Consciousness Analysis", in Proceedings of IEEE EMBC Conference, pp. 810-183, 2011

[12] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy: A tool for complexity analysis of multichannel data", Physical Review E, vol. 84, no. 6, pp. 061918-1 061918-10, 2011

[13] M. U. Ahmed, N. Rehman, D. Looney, T. Rutkowski, and D. P. Mandic, "Dynamical Complexity of Human Responses: A Multivariate Data-Adaptive Framework", Bulletin of the Polish Academy of Sciences, in print, 2012

[14] D. Looney, M. U. Ahmed and D. P. Mandic, "Human Centred Multivariate Complexity Analysis", Natural Intelligence Magazine, vol 1, no. 3, pp. 40-42, 2012

[15] http://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html

[16] http://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html