Computers in Human Behavior 27 (2011) 1512–1518

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Interactive component extraction from fEEG, fNIRS and peripheral biosignals for affective brain–machine interfacing paradigms

Tomasz M. Rutkowski [a,*], Toshihisa Tanaka [b,a], Andrzej Cichocki [a], Donna Erickson [c,a], Jianting Cao [d,a], Danilo P. Mandic [e]

[a] Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi 351-0198, Japan
[b] Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, 2-24-16 Nakacho, Koganei-shi, Tokyo 184-8588, Japan
[c] Showa Music University, 1-11-1 Kamiasao, Asao-ku, Kawasaki-shi, Kanagawa 215-8558, Japan
[d] Saitama Institute of Technology, Saitama 369-0293, Japan
[e] Communication and Signal Processing Research Group, Department of Electrical and Electronic Engineering, Imperial College, Exhibition Road, London SW7 2BT, United Kingdom

ARTICLE INFO

Article history:
Available online 3 November 2010

Keywords:
Brain/human–computer interaction
EEG
Intelligent brain signal processing
EMD application to EEG
Interactive brain–computer interfacing
Affective human/brain–computer interfacing paradigm

ABSTRACT

This paper investigates whether some well understood principles of human behavioral analysis can be used to design novel paradigms for affective brain–computer/machine interfaces. This is achieved by using the visual, audio, and audiovisual stimuli representing human emotions. The analysis of brain responses to such stimuli involves several challenges related to the conditioning of brain electrical responses, extraction of the responses to stimuli and mutual information between the several physiological recording modalities used. This is achieved in the time–frequency domain, using multichannel empirical mode decomposition (EMD), which proves very accurate in the joint analysis of neurophysiological and peripheral body signals. Our results indicate the usefulness of such an approach and confirm the possibility of using affective brain–computer/machine interfaces.

1. Introduction

“Human factors” and “human behavioral” aspects in design of contemporary interactive communication systems play an important role, especially when efficiency is paramount, as shown by Rutkowski, Mandic, and Barros (2007). Human–computer/machine-interfacing (HCI/HMI) and a more recent field of brain–machine/computer-interfacing (BCI/BMI) (as reviewed by Cichocki et al. (2008)), are emerging topics of research, lending themselves to more smooth or natural interaction between humans and machines/computers. While HCI/HMI interfaces rely on peripheral or more generally muscle movements learned or adopted by users, BCI/BMI technologies aim to interpret brain activity which precedes real movements or is related to movements/actions planning (Cichocki et al., 2008; Wolpaw & McFarland, 2004; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002).

Already established neuroscience tools such as functional electroencephalography (fEEG) and functional near infrared spectroscopy (fNIRS) correlate conscious and affective experiences with electric field activity and oxygenation changes localized in cortical areas of the brain. Additional peripheral body measurements such as skin conductance, heart-rate, breath and pulse variability, as well facial muscle and eye-movement characteristics also correlate to emotional arousal (Rutkowski et al., 2007, Rutkowski, Cichocki, Ralescu, & Mandic, 2008b). These physically based measures provide an objective way to explore the realm of perception, experience, mind and emotional processes estimate in real-time from human subjects exposed to external stimuli. The multimodal stimuli can be presented in the form of mythology, stories, and multimedia through the use of imagination, images, music, sounds, and movies. All these influence the mind to evoke a wide range of emotions (Rutkowski et al., 2008b). Emotional or more generally affective stimuli are chosen by the authors due to their importance in decision making process in human brains as discussed extensively by Lehrer (2009).

Today’s interactive media such as video games provide a highly interactive platform to test how users interact with the environment based upon their unique experiences and anatomical structure of prefrontal cortices. Interactive multimedia combined with neuro- or biofeedback provide a unique platform for conducting objective investigations into the mind–body relationship; in connection with interactive communication (Rutkowski et al., 2007) paradigms these can be further utilized in brain–machine-interfacing technologies.

* Corresponding author. Tel.: +81 484679666; fax: +81 4679694.
E-mail addresses: tomek@brain.riken.jp (T.M. Rutkowski), tanakat@cc.tuat.ac.jp (T. Tanaka), a.cichocki@riken.jp (A. Cichocki), ericksondonna2000@gmail.com (D. Erickson), cao@sit.jp (J. Cao), d.mandic@imperial.ac.uk (D.P. Mandic).

0747-5632/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2010.10.016

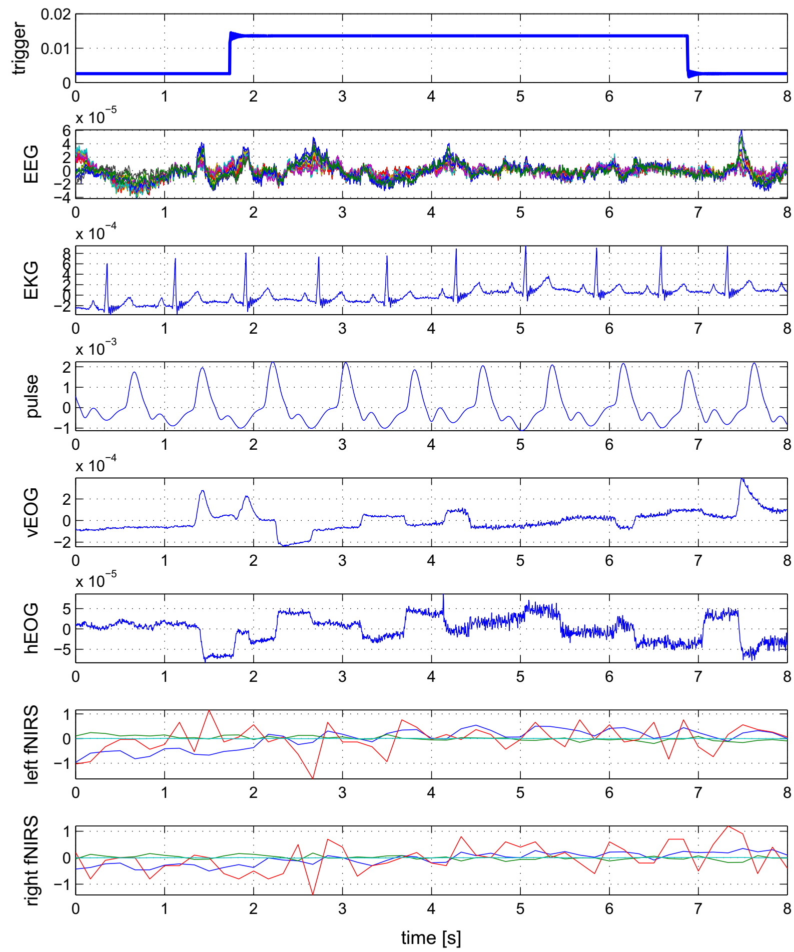

Fig. 1. Multimodal bio-signals recorded from subject's body surface. The top panel presents stimulus onset and offset times. The second from the top panel presents 16 fEEG channels plotted together, while next two panels depict ECG and pulse oximetry time series. The two following panels labeled vEOG and hEOG show vertical and horizontal eye movements, respectively. The two bottom panels depict fNIRS recordings from left and right frontal cortices. All signals were recorded synchronously with g.USBamp and NIRO-200 devices connected to a single workstation running Matlab.

Recent advances in BCI/BMI technologies also reveal a need to search for new and more challenging paradigms which would allow more natural interaction of humans and machines by exploiting the information in these new communication channels (Cichocki et al., 2008).

There are two recognized general classes of BCI/BMI paradigms: those related to external environment stimuli and the other which are completely independent from environmental stimulation and relay only on internal (imaginary) brain activity managed by the users will. The second class of imaginary paradigms is usually more difficult for non-trained subjects, since they require learned brain activity patterns to be captured by non-invasive brain activity methods such as fEEG and fNIRS. In this paper, we focus on the first class of dependent and stimuli driven paradigms, with an interactivity concept involved, utilizing emotional empathy paradigms (Mehrabian & Epstein, 1972). We provide insights into mind–body relationship and show that in general brain activity related signals are correlated with peripheral electrophysiological and physiological responses.

2. Methods

In order to evaluate mind–body interactions and illustrate the possibility to recognize patterns of reactions to presented multimodal emotional stimuli, the following experimental procedure is proposed. The recording experiments combining fEEG, fNIRS and peripheral electrophysiological signals were conducted at the Advanced Brain Signal Processing Laboratory of the RIKEN Brain Science Institute, Wakoshi, Japan. The experimental procedure followed the guidelines for experiments with human subject of institute's ethical committee. Brain and body electrophysiological responses were recorded using two synchronized g.USBamp biosignal data acquisition systems with 16 fEEG electrodes placed over frontal, temporal and parietal lobes; two channels of vertical and horizontal EOG; a single ECG channel; and pulse. Additionally two frontal fNIRS channels were recorded synchronously with NIRO-200 cerebral oxygenation recorder. An example of such multimodal recording is shown in Fig. 1.

The subjects were given audio-only and video-only affective presentations from the emotional utterances corpus (Baron-Cohen, 2004) as performed by five British English professional actors. Both the video and audio presentations portrayed affective expressions of six basic emotions. The video-only presentations involved short (2–5 s long) movies; the audio-only involved short (also 2–5 s long) sentences. After attaching the monitoring electrodes, the subjects were instructed to look at a white cross mark on the computer screen and to try not to blink or move in order to minimize muscular noise. The subjects were instructed to answer a question on the screen after the audio or visual presentation which emotion did they perceive. The purpose of these questions was to focus subjects attention on the task and to give them a period of relaxation time, and their answers were not analyzed. The main goal of the experiment was a search for interactive ("emotional synchrony") responses captured within neurophysiological and peripheral electrophysiological signals carrying very short emotional empathy

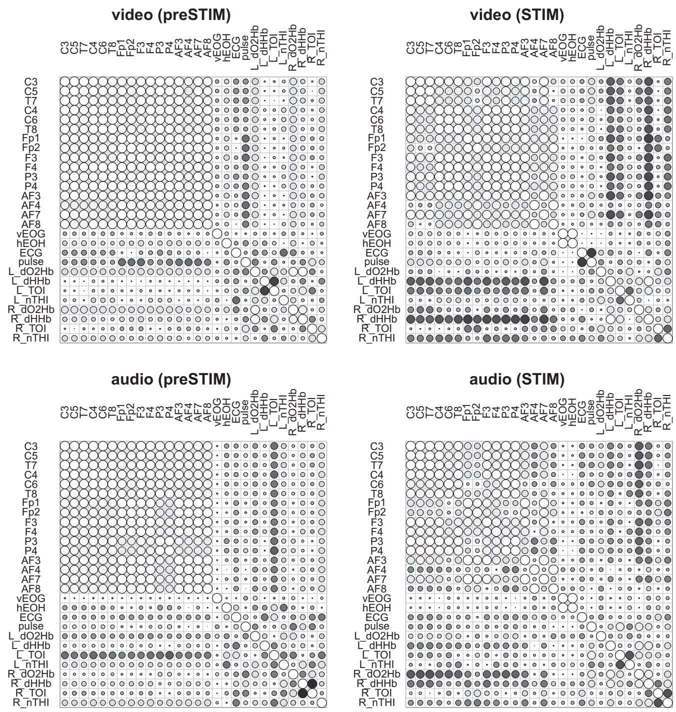

Fig. 2. Emerging patterns of fEEG, fNIRS and peripheral bio-signals for pre-stimulus and stimulus conditions compared in the form of multimodal correlation data analysis results for a subject stimulated with *angry* emotional display in visual (top panels) and auditory (bottom panels). The channels numbers are as follows: 1–16 = fEEG; 17 = ECG; 18 = horizontal EOG; 19 = vertical EOG; 20 = pulse (*pulseoximetry*); 21–24 = left forehead fNIRS: *L–dO2Hb*, *L–dHHb*, *L-TOI*, and *L–nTHI*, respectively; 25–28 = right forehead fNIRS: *R–dO2Hb*, *R–dHHb*, *R–TOI*, and *R–nTHI*, respectively. The sizes of correlation matrix rings visualize the coefficients values, where black color = -1 and white = +1. In this result the correlation patterns change within fEEG (1–16) channels as well fEEG/fNIRS (1–16)/(21–26) in both auditory and visual condition identifies subject's interactive response.

signatures. A concept of empathy is characterized as a capability to share ones feelings and understand another's emotion and feelings and it was shown previously by the authors that empathy response could be recognized and classified from the fEEG responses only (Rutkowski et al., 2008b).

2.1. Multimodal biosginals preprocessing

The multimodal fEEG, fNIRS, EOG, ECG and pulse signals (see Fig. 1) had to be first preconditioned, due to their different sampling frequencies and dynamics. In order to obtain common coherent interactive responses carrying empathy responses an approach described in Rutkowski et al. (2009) was utilized. All signals were first decomposed using empirical mode decomposition (EMD), and later clusters of similar components in Huang–Hilbert spectral domain were obtained. This method allowed for identification of those components within each channel which exhibited similar spectral patterns across all data, as well as those synchronized with onsets and offsets of the stimuli as shown in the top panel of Fig. 1.

2.1.1. Application of EMD to multichannel EEG

EMD utilizes empirical knowledge of oscillations intrinsic to a signal in order to represent them as a superposition of components, defined as *intrinsic mode functions* (IMF), with well defined instantaneous frequencies. In order to obtain an IMF from a single channel of a biosignal recording it is necessary to remove local riding waves and asymmetries (originating usually from non-body-related sources), which are estimated from local envelope of minima and maxima of the waveform being processed. Search for IMFs corresponds thus to separation of band-limited and semi-orthogonal components from the an analyzed signal. It also corresponds to eliminating riding-waves, which ensures that the IMF will have no fluctuations caused by an asymmetric waveform. In each decomposition cycle, the IMF is defined by zero crossings and involves only one mode of oscillation (thus complex waves are not allowed). Notice that IMF is not limited to be a narrow band signal, as it would be in case of contemporary time–frequency analysis methods (Fourier analysis, wavelets, etc.). In fact, an IMF can be both amplitude and frequency modulated simultaneously, as well as non-stationary or non-linear, which is a strongest point of this technique applied for as complex signals as EEG, EMG, EOG, ECG, fNIRS, etc.

EMD decomposes the biosignals within each channel modality into IMFs (Huang et al., 1998). The method allows to represent them in form of "oscillatory modes" which satisfy the following conditions:

(i) The number of extrema and the number of zero crossings of each IMF should be either equal or differ at most by one.
(ii) At any point, the mean value of the envelopes defined by the local maxima and the local minima should be zero.Since IMF represents an oscillatory mode within a signal; its periods, which are defined by zero crossings, correspond to the only *single* mode of oscillation which is a frequency-band-limited activity preferably originating from a single source in human body or brain. Both the amplitude and frequency of this oscillation may fluctuate over time, in other words, the oscillation is not necessarily stationary nor narrow-band.

The process of extracting an IMF from a signal $x(t)$ is called "sifting" (Huang et al., 1998) and it consists of the following steps:

1. First determine the local maxima and minima of the signal $x(t)$.
2. Next generate the upper and lower signal envelopes by connecting the local maxima and minima, respectively, utilizing an interpolation method (e.g., linear, spline, piece-wise spline (Huang et al., 1998, Rutkowski, Cichocki, & Mandic, 2008a)).

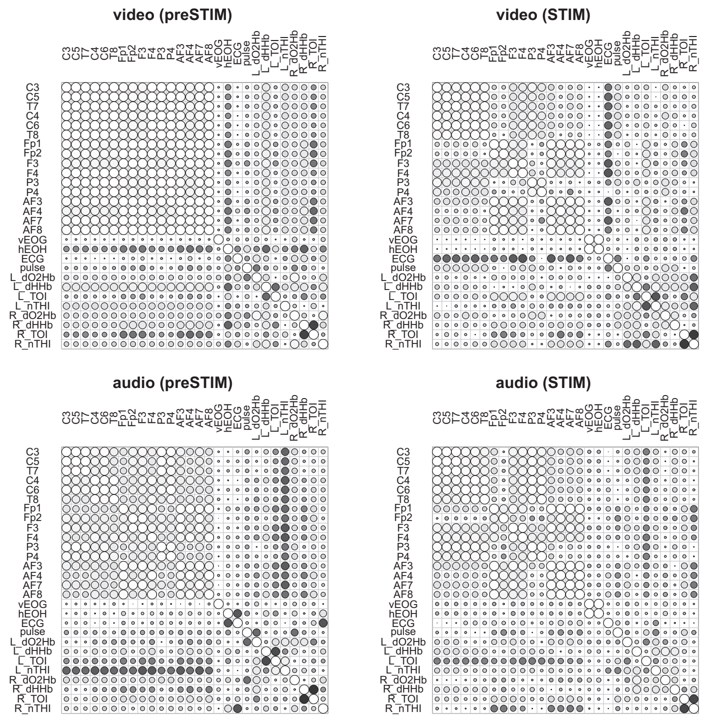

Fig. 3. Emerging patterns of fEEG, fNIRS and peripheral bio-signals for pre-stimulus and stimulus conditions compared in the form of multimodal correlation data analysis results for a subject stimulated with *neutral* emotional display in visual (top panels) and auditory (bottom panels). The channels numbers are as follows: 1–16 = fEEG; 17 = ECG; 18 = horizontal EOG; 19 = vertical EOG; 20 = pulse (*pulseoximetry*); 21–24 = left forehead fNIRS: *L–dO2Hb*, *L–dHHb*, *L–TOI*, and *L–nTHI*, respectively; 25–28 = right forehead fNIRS: *R–dO2Hb*, *R–dHHb*, *R–TOI*, and *R–nTHI*, respectively. The sizes of correlation matrix rings visualize the coefficients values, where black color = -1 and white = +1. In this result the correlation patterns change within fEEG channels (1–16) in both auditory and visual condition identifies subject's interactive response.

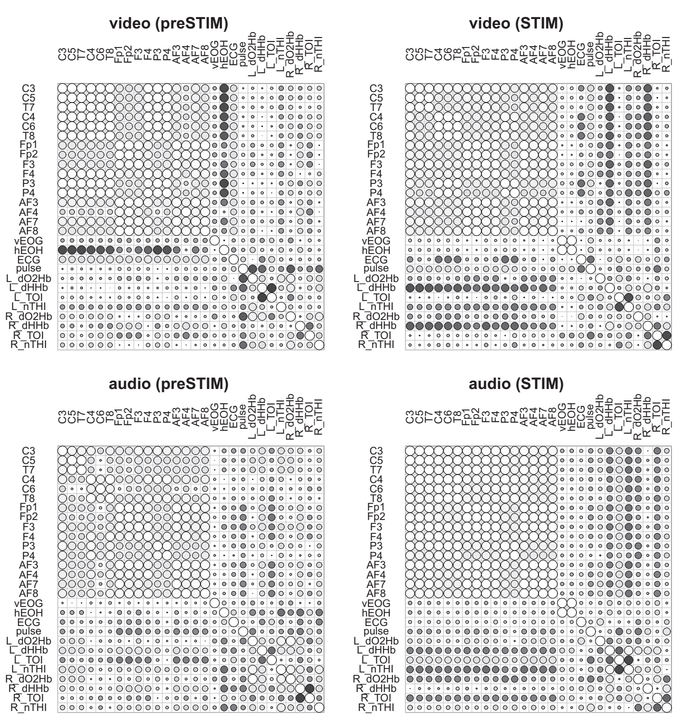

Fig. 4. Emerging patterns of fEEG, fNIRS and peripheral bio-signals for pre-stimulus and stimulus conditions compared in the form of multimodal correlation data analysis results for a subject stimulated with *positive* emotional display in visual (top panels) and auditory (bottom panels). The channels numbers are as follows: 1–16 = fEEG; 17 = ECG; 18 = horizontal EOG; 19 = vertical EOG; 20 = pulse (*pulseoximetry*); 21–24 = left forehead fNIRS: *L–dO2Hb*, *L–dHHb*, *L–TOI*, and *L–nTHI*, respectively; 25–28 = right forehead fNIRS: *R–dO2Hb*, *R–dHHb*, *R–TOI*, and *R–nTHI*, respectively. The sizes of correlation matrix rings visualize the coefficients values, where black color = -1 and white = +1. In this result the correlation patterns change of fEEG/fNIRS (1–16)/(21–26) in both auditory and visual condition indicates subject's interactive response.

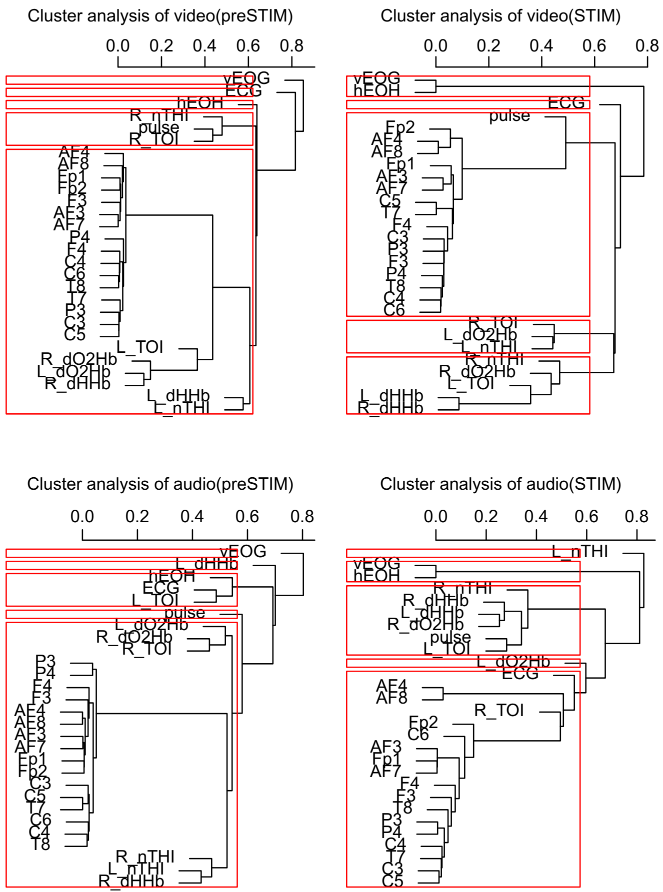

Fig. 5. Hierarchical clustering result of multimodal biosignal correlation patterns visualized in Fig. 2 with five major clusters depicted in red. The so obtained clusters for pre-stimuli and stimuli conditions of *angry* emotional display in visual (top panels) and auditory (bottom panels) present similar rearrangement with very interesting transition of both EOG channels into a single cluster and separation of fEEG and fNIRS modalities. In both stimuli cases also *ECG* modality became separated into a single cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

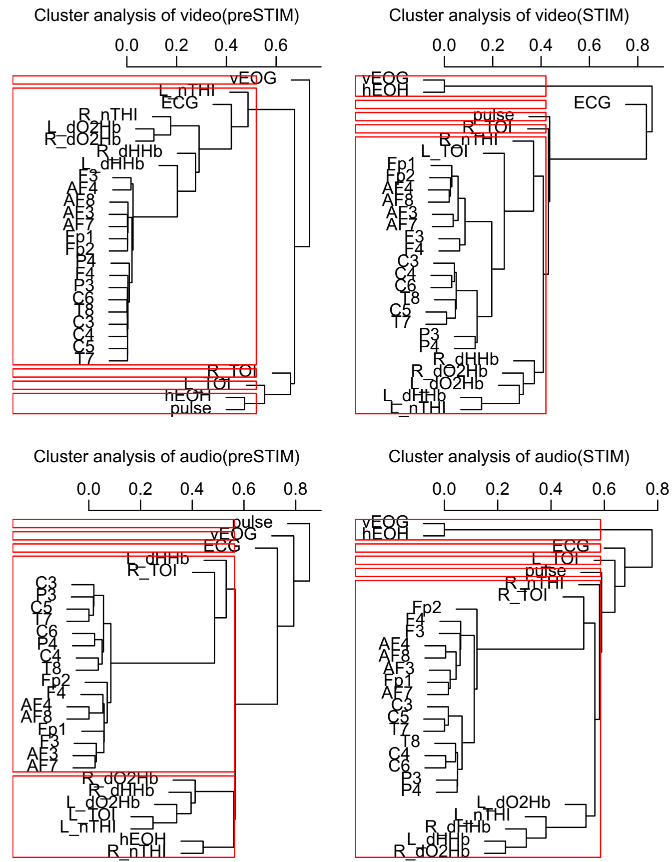

Fig. 6. Hierarchical clustering result of multimodal biosignal correlation patterns visualized in Fig. 3 with five major clusters depicted in red. The so obtained clusters for pre-stimuli and stimuli conditions of *neutral* emotional display in visual (top panels) and auditory (bottom panels) present opposite rearrangement (compare Fig. 5) of fEEG and fNIRS modalities into a single cluster, but with a similar transition of both EOG channels into a single cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. After that calculate a local mean $m_1(t)$, by averaging the upper and lower signal envelope.
4. Finally subtract the local mean from the data:

$$h_1(t) = x(t) - m_1(t). \tag{1}$$

Ideally, $h_1(t)$ should satisfy the above criteria 1.–4. of an IMF estimation, however, typically this procedure needs to be repeated until the first IMF is successfully extracted. In order to obtain the second IMF the sifting process is applied to the residue $\varepsilon_1(t) = x(t) - \mathrm{IMF}_1(t)$, obtained by subtracting the first IMF from $x(t)$; the third IMF is in turn extracted from the residue $\varepsilon_2(t)$ and so on. The decomposition is finished when the two consecutive sifting results are similar and no more IMF could be found. The empirical mode decomposition of the signal $x(t)$ may be written as a summation of n estimated IMFs and a residuum $\varepsilon_n(t)$ which is either a mean trend originating from amplifier's baseline drift or a constant:

$$x(t) = \sum_{k=1}^{n} \mathrm{IMF}_k(t) + \varepsilon_n(t). \tag{2}$$

2.2. Correlation patterns analysis

The preprocessed multimodal neurophysiological and peripheral electrophysiological signals carrying only components exposing synchrony with the emotional stimuli presented to the subjects (a trigger signal marking stimuli onsets and offsets as in top panel of Fig. 1) can now be analyzed for their cross-correlations patterns during subject interactions with them. This is visualized in the form of scatter plots of pairwise Pearson's correlation coefficients matrices as in Figs. 2–4, for emotional stimuli displays of *angry*, *neutral* and *positive* in visual and auditory domains, presented to the subjects. The Pearson's correlation coefficient values in the figures are visualized in the form of rings whose sizes depict the values, whereby the black color indicates negative and white positive ones. For each visual and auditory stimuli conditions the cases of pre-stimuli (*preSTIM*) and emotional display presentation (*STIM*) multimodal biosignal windows are presented. Very interesting and stable within different emotional display changes in brain and peripheral biosignals patterns can be observed.

2.3. Clustering of preprocessed biosignals

In order to further analyze and visualize differences in responses three emotional displays within all preprocessed biosignals we evaluate their similarity across the channels for both pre-stimuli and stimuli conditions. A distance measure chosen was (1 − *correlation*) based on Pearson's correlations coefficients between vectors being evaluated similarly as Section 2.2. We discussed several distance measures possible to utilize in our previous work (Rutkowski, Toshihisa, Cichocki, & Mandic, 2008c) and the above mentioned (1 − *correlation*) was identified as the

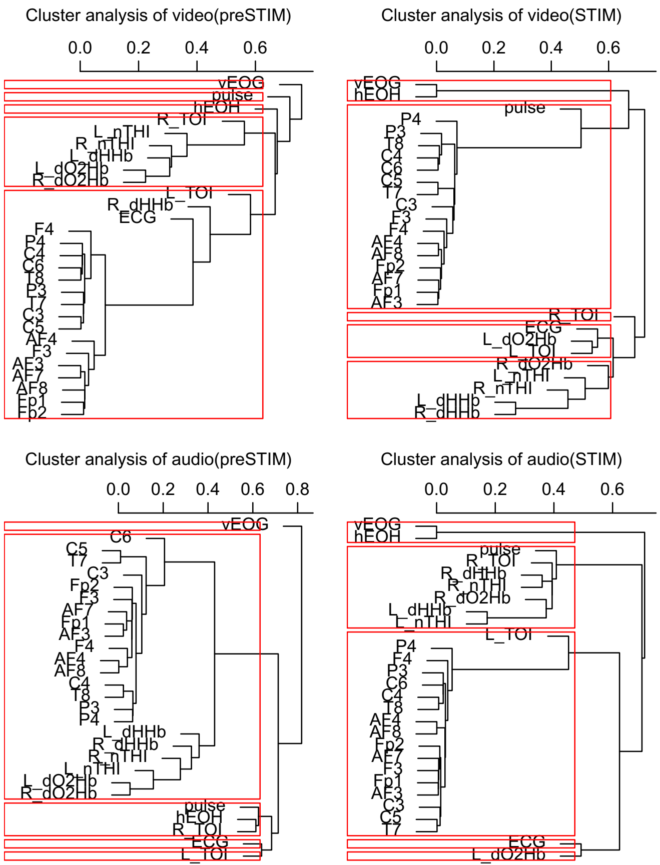

Fig. 7. Hierarchical clustering result of multimodal biosignal correlation patterns visualized in Fig. 4 with five major clusters depicted in red. The so obtained clusters for pre-stimuli and stimuli conditions of *positive* emotional display in visual (top panels) and auditory (bottom panels) present similar rearrangement with very interesting transition of both EOG channels into a single cluster (similar as in both cases in Figs. 5 and 6) and separation of fEEG into a single and fNIRS into two clusters modalities. In both stimuli cases also *ECG* modality became separated into a single cluster similar as in previous case visualized in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

most suitable for multimodal biosignals recorded during interaction with affective stimuli.

The distances among preprocessed signals were evaluated in form of a hierarchical cluster analysis (HC). In this approach initially each vector representing each modality channel was assigned to its own cluster (*bottom-up-strategy*) and then the algorithm proceeded iteratively, at each stage joining the two most similar clusters. Such procedure continued until there was just a single cluster. At each stage distances between clusters were recomputed by a Lance–Williams dissimilarity update formula with a single linkage method clustering method. The single linkage method is closely related to the minimal spanning tree concept and it adopts a "friends of friends" strategy for clustering (Murtagh, 1985).

The clusters of components obtained with the above two strategies are presented in Figs. 5–7 for affective responses to *angry*, *neutral* and *positive* emotional stimuli. In all cases very interesting and stable among visual and auditory modalities clustering transitions were observed allowing to identify with which visual or auditory emotional stimuli subject was interacting.

For affective responses to *angry* emotional display stimuli (see Fig. 5) a rearrangement with very interesting transition of both EOG channels into a single cluster and separation of fEEG and fNIRS modalities was observed. In both stimuli cases also *ECG* modality became separated into a single cluster exposing "desynchrony" of heart activity for remaining biosignal modalities.

In case of *neutral* emotional display an opposite rearrangement (compare *angry* example in Fig. 5) of fEEG and fNIRS modalities into a single cluster, but with a similar transition of both EOG channels into a single cluster was observed.

The third case of 'positive" emotional displays presents similar rearrangement with very interesting transition of both EOG channels into a single cluster (similar as in both *angry* and *neutral* cases in Figs. 5 and 6) and separation of fEEG into a single and fNIRS into two clusters modalities. In both stimuli cases also *ECG* modality became separated into a single cluster similar as in previous case visualized in Fig. 6.

3. Conclusions

We have observed changes in correlation patterns during interactive presentation of three different emotional stimuli in visual and auditory domains as depicted in Figs. 2–4. The internal fEEG, fEEG/fNIRS as well fEEG/fNIRS/biosignals correlation patterns showed significant differences for the three experimental conditions. Further conducted hierarchical cluster analysis of so obtained Pearson's coefficients utilized as distance measures confirmed existence of different response patterns to various emotional stimuli in brain and peripheral body measures. This opens a possibility to utilize emotional responses in human–computer/machine interaction in real world conditions.

The observed mind–body correlations pattern changes for prestimuli and stimuli conditions show potential possibility in human–computer/machine interactions design based on behavioral based interfaces using an online feedback from the users connected/wired as in classical BCI/BMI applications. Smooth and natural human–machines interactions should adopt to users mental/emotional states.

We have shown than that interactive empathy responses to emotional stimuli in auditory and visual domains are good candidates for their utilization in intelligent computing applications such as BCI/BMI since it is possible to discriminate the response patterns from neurophysiological signals (fEEG and fNIRS) together with periphery electrophysiological ones (ECG, EOG, pulse).

Acknowledgements

The research presented in this paper was supported in part by the Japan Society for the Promotion of Science Grant No. 21360179 and by the AOARD and Royal Society International Joint Project grants.

References

Baron-Cohen, S. (2004). *Mind reading – The interactive guide to emotions*. London, UK: Jessica Kingsley Publishers.

Cichocki, A., Washizawa, Y., Rutkowski, T., Bakardjian, H., Phan, A.-H., Choi, S., et al. (2008). Noninvasive BCIs: Multiway signal-processing array decompositions. *Computer, 41*(10), 34–42.

Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., et al. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 454*(1971), 903–995.

Lehrer, J. (2009). *How we decide*. Boston & New York: Houghton Mifflin Harcourt Publishing Company.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*(4), 525–543.

Murtagh, F. (1985). *Multidimensional clustering algorithms*. Physica-Verlag.

Rutkowski, T., Mandic, D., & Barros, A. (2007). A multimodal approach to communicative interactivity classification. *The Journal of VLSI Signal Processing, 49*(2), 317–328.

Rutkowski, T. M., Cichocki, A., & Mandic, D. P. (2008a). Signal processing techniques for knowledge extraction and information fusion. Information technology: Transmission, processing and storage. In *EMDsonic – An empirical mode decomposition based brain signal activity sonification approach* (pp. 261–274). Springer. Ch..

Rutkowski, T. M., Cichocki, A., Ralescu, A. L., & Mandic, D. P. (2008b). Emotional states estimation from multichannel EEG maps. In R. Wang, F. Gu, & E. Shen (Eds.), *Advances in cognitive neurodynamics ICCN 2007 proceedings of the international conference on cognitive neurodynamics. Neuroscience* (pp. 695–698). Berlin & Heidelberg: Springer.

Rutkowski, T. M., Toshihisa, T., Cichocki, A., & Mandic, D. P. (2008c). Clustering EMD components for muscular interference separation from EEG – A time/frequency approach with different distance measures. In *Proceedings of 23rd SIP symposium* (pp. 52–57). Kanazawa: IEICE.

Rutkowski, T. M., Cichocki, A., Tanaka, T., Mandic, D. P., Cao, J., & Ralescu, A. L. (2009). Multichannel spectral pattern separation – An EEG processing application. In *Proceedings of the 2009 IEEE international conference on acoustics, speech, and signal processing (ICASSP2009)* (pp. 373–376). Taipei, Taiwan: IEEE Press.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology, 113*, 767–791.

Wolpaw, J. R., & McFarland, D. J. (2004). Control of a two-dimensional movement signal by a noninvasive brain–computer interface in humans. *Proceedings of National Academy of Sciences of the United States America, 101*(51), 17849–17854.