Continuous Recognition of Affective States by Functional Near Infrared Spectroscopy Signals

Dominic Heger, Reinhard Mutter, Christian Herff, Felix Putze, and Tanja Schultz
Karlsruhe Institute of Technology (KIT)
Adenauerring 4, 76131 Karlsruhe, Germany
Email: dominic.heger@kit.edu

***Abstract*—Functional near infrared spectroscopy (fNIRS) is becoming more and more popular as an innovative imaging modality for brain computer interfaces. A continuous (i.e. asynchronous) affective state monitoring system using fNIRS signals would be highly relevant for numerous disciplines, including adaptive user interfaces, entertainment, biofeedback, and medical applications. However, only stimulus-locked emotion recognition systems have been proposed by now. fNRIS signals of eight subjects at eight prefrontal locations have been recorded in response to three different classes of affect induction by emotional audio-visual stimuli and a neutral class. Our system evaluates short windows of five seconds length to continuously recognize affective states. We analyze hemodynamic responses, present a careful evaluation of binary classification tasks and investigate classification accuracies over the time.**

I. Introduction

Emotions are psycho-physiological processes that play a key role in human life. They are an essential part of social interaction and can have strong regulatory influence on many important conscious and unconscious aspects of human behavior, including influences on perception, learning and decision making. Since the 2000s, affective computing [1] has grown to a highly vivid area of research. Today, intelligent machines that sense human emotions in order to act in a natural, human-like and empathic way, can still only be found in science fiction movies. However, a few attempts have successfully shown that systems adapting to cognitive or affective user states can be advantageous and preferred by users (e.g. [2], [3]). In addition to that, affective computing technology has strongly contributed to multiple disciplines besides the field of human-machine interaction. For example, emotion recognition systems have been used in medical applications, such as autism spectrum disorders therapy (e.g. [4]). In the future, numerous application could benefit from a possibility to continuously monitor a person's affective states, including adaptive user interfaces, biofeedback, and the assessment of the quality of life of patients in clinical settings (bed-site monitoring of emotions).

Affective states manifest themselves in various biosignals. However, they originate in different brain circuits that relate to emotions. Besides subcortical areas, such as parts of the limbic system (e.g. amygdala and anterior cinguate cortex) [5], [6], [7], the essential role of the prefrontal cortex (PFC) in emotion induction and regulation has been shown (e.g. [8]).
For several years, Brain Computer Interfaces have successfully been developed that provide strongly motor impaired patients a way of communication and control of computing devices. More recently, so-called passive BCIs [9] gain rising attention. Instead of using an explicit control intention, they aim at passively observing information about cognitive or affective mental states from the users' brain activity. Several passive BCI systems for monitoring of user states such as attention, workload, and also affective states have been proposed. Most of these systems use EEG as input modality [10]. Since the last few years, fNIRS is becoming more and more popular as a non-invasive measurement technique for BCIs.
fNIRS is an imaging modality detecting changes in regional cerebral blood flow. According to the blood oxygen level-dependent (BOLD) effect, oxygenated (HbO_2) and deoxygenated (HbR) hemoglobin are functional indicators for brain activity. fNIRS, exploits the fact that HbO_2 and HbR have different absorption rates for light in the near infrared part of the spectrum. As near infrared light disperses through biological tissue, but is absorbed by hemoglobin, light sources and detector optodes placed on the subjects' scalp can be used for functional imaging. Using the modified Beer-Lambert Law [11] changes in the cerebral blood flow, and thereby brain activity, can be estimated from the changes in light intensities. Measurement positions arise roughly in the middle between transmitter and detector in a depth of half the source detector distance. Typically, HbO_2 levels in a cortical area rise with brain activity and peek approximately 5 to 10 seconds after the beginning of activation, HbR levels should fall in the same intervals. Compared to fMRI, which also measures the BOLD response, fNIRS is comparably cheap, portable and does not confine the subjects. In comparison to EEG, fNIRS is not susceptible to electrical artifacts from environmental and physiological sources. Furthermore, no conductive gel needs to be used and frontal fNIRS recordings, where measurements are not obstructed by hair, have very short setup times (about one minute).

In this paper, we explore advantages and limitations of a system for the continuous (i.e. asynchronous) decoding of affective states from fNIRS signals recorded at prefrontal locations in response to audio-visual stimulation. To the best of our knowledge, fNIRS studies on emotion recognition have only investigated event-related responses to emotional stimuli using stimulus locked evaluations, before. However, such setups are primarily restricted to laboratory conditions. In most realistic scenarios the point of time of an emotion triggering event and the time span of an affective state are unknown. Our system classifies simple features extracted from short slices of the fNIRS signal stream over time and continuously outputs estimates of the affective user state.

A. Related Work

Numerous pattern recognition systems for the automatic classification of emotions from speech, visual, or biophysiological data have been proposed (see e.g. [12], [13] for review). They typically follow a pattern recognition approach, where features are extracted from preprocessed input signals and classified by machine learning methods. For training data collection, mostly controlled laboratory settings are used. The perception of emotional stimuli has been shown to be an effective method for emotion induction in such experiments. Standardized picture sets and sound samples for emotion induction are available. In numerous studies emotional stimuli based on the International Affective Picture System (IAPS) [14] and International Affective Digital Sounds (IADS) [15] have been used. To quantitatively describe emotions, discrete classes of emotions (e.g. Ekman's basic emotions [16]) or dimensional emotion models (e.g. Russell's Pleasure-Arousal-Dominance scale [17]), are most widely used. IAPS and IADS have been rated using the Self Assessment Manikins in the dimensions pleasure, arousal and dominance by large person groups [18], [15].

Several studies investigated fNIRS signals in the context of emotions and showed significant activations of the prefrontal cortex in response to emotionally loaded pictures. Herrmann et al. [19] showed that the prefrontal cortex is activated in response to emotional induction by pictures and facial expressions using fNIRS and investigated these effects with regard to task requirements. Hoshi et al. [20] found that most subjects showed decreases in HbO_2 in multiple channels during the presentation of pleasant pictures. Half of the subjects showed significant increases of HbO_2 during unpleasant pictures. Katsuhiko et al. [21] reported activations of HbO_2 in right frontal regions during anticipatory anxiety. Further studies reported gender differences in prefrontal activations during emotion induction [22], [23] and investigated differences in emotion induction and emotion regulation [24].

In the last few years, fNIRS based Brain Computer Interfaces have been developed that use different task paradigms, including motor imagery, mental arithmetics, mental workload and speech and language (e.g. [25], [26], [27], [28], [29]). To the best of our knowledge Tai and Chau [30] were the first who investigated single-trial recognition of an emotion task. They recorded 16 channels fNIRS from the prefrontal cortex in response to affective induction using emotional images. Their system applied a genetic algorithm to find the optimal analysis interval length, feature set, and classifier. They report very high recognition rates of 75-97%, however, they only investigated discriminating baseline periods from emotion induction periods and not different emotion classes. Hosseini et al. [31] decoded if subjects like or dislike different visually presented objects using fNIRS signals recorded from the anterior frontal cortex. Moghimi et al. [32] showed that emotional responses to music can automatically be classified using prefrontal fNIRS signals. They discriminated emotional music from brown noise and emotional valence (positive versus negative) with average recognition accuracies above 70%. Asano et al. [33] induced comfortable and uncomfortable affective states by different sounds (music and scratch sounds). They measured 32 channels of frontal fNIRS and discriminated two affective states using Bayesian nets with an average accuracy of 67%.

II. Material and Experimental Design

A. Participants

8 healthy male volunteers participated in the experiment (mean age 27.6±5.2 years). All subjects had normal or corrected to normal vision and normal hearing. None of the subjects had a history of brain injury, cardiovascular disease, drugs, or psychiatric conditions. Subjects have been instructed to the experiment and provided informed consent before the start of the recordings.

B. Stimuli

We composed a slideshow for audio-visual emotional induction containing stimuli from the International Affective Picture System (IAPS) [14] and International Affective Digital Sounds (IADS) [15]. Pictures and sounds with extreme ratings in valence and arousal according to the self assessment ratings in [18], [15] were selected and categorized into the following three emotion classes:

- **VA** - Maximum valence and maximum arousal: This class contained e.g. pictures of female erotica and exciting sports, rhythmic music and shouts of joy.
- **Va** - Maximum valence and minimum arousal: This class contained e.g. pictures of flowers and animals, calm classical music and environmental sounds.
- **vA** - Maximum arousal and minimum valence: This class contained e.g. pictures of multilated persons and threatening situations and sounds from air raid, and screaming persons.

The picture and sound items[1] were selected to achieve a strong emotional stimulation and discriminability between the classes. Similar slideshows for audio-visual emotion induction have successfully been used in one of our previous studies [34], where subjects' self-assessments rated the stimuli as having a strong emotional impact corresponding to the emotion classes in terms of valence and arousal. Table I summarizes average subjective ratings of the pictures and sounds used in the experiment.

C. Experimental procedure

Affective pictures were presented on a computer screen located in about 0.5 meters distance, while subjects were

[1]The following IAPS picture items have been used:
VA: 4001, 4002, 4003, 4006, 4141, 4142, 4150, 4180, 4220, 4225, 4232, 4235, 4250, 4255, 4274, 4275, 4279, 4290, 4300, 4310, 4311, 4607, 4608, 4651, 4652, 4659, 4660, 4664, 4670, 4683, 4694, 4695, 4800, 8030, 8080, 8185, 8400, 8501; **Va**: 1440, 1441, 1460, 1560, 1590, 1600, 1601, 1602, 1603, 1604, 1610, 1620, 1640, 1670, 1900, 2000, 2050, 2070, 2170, 2260, 2304, 2341, 2360, 2370, 2501, 5000, 5010, 5020, 5030, 5200, 5551, 5720, 5760, 5800, 5891, 7080, 7325, 7545, 7900; **vA**: 3000, 3010, 3015, 3030, 3051, 3053, 3060, 3062, 3063, 3068, 3069, 3071, 3080, 3101, 3102, 3110, 3120, 3160, 3168, 3170, 3181, 3250, 3261, 3266, 3301, 3400, 3500, 3530, 6230, 6260, 6313, 6510, 6540, 9400, 9405, 9410, 9433, 9570, 9635.1, 9810
The following IADS sound items have been used:
VA: 201, 202, 311, 215, 815, 716, 360, 200, 366, 717, 817, 205, 352, 367, 415, 204, 353, 216, 110, 355; **Va**: 112, 150, 151, 171, 172, 230, 262, 270, 370, 374, 377, 602, 705, 721, 725, 726, 809, 810, 811, 812; **vA**: 115, 260, 275, 276, 277, 278, 279, 284, 285, 286, 290, 292, 420, 422, 424, 600, 624, 709, 711, 712

TABLE I. SUBJECTIVE RATINGS ACCORDING TO [18] AND [15] OF THE SELECTED PICTURES AND SOUNDS OF THE THREE EMOTION CLASSES. MEANS (STANDARD DEVIATIONS) ON A 9 POINT SCALE.

Emotion Class	IAPS		IADS	
	Valence	Arousal	Valence	Arousal
VA	7.48 (1.42)	6.94 (1.88)	7.13 (1.79)	6.95 (1.80)
vA	2.25 (1.49)	6.17 (2.31)	2.41 (1.61)	7.29 (1.81)
Va	6.71 (1.58)	3.49 (2.10)	6.31 (1.66)	4.14 (2.06)

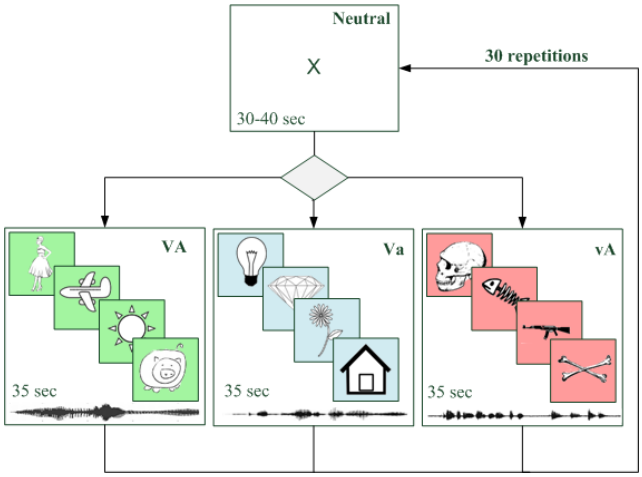

Fig. 1. Experimental protocol for data collection.

sitting in a comfortable office chair. Sounds were played using headphones at a moderate volume that was set identical for all subjects. Each subject was instructed just to watch the slideshow and to avoid unnecessary movements.
During one recording session subjects attended 30 emotion induction blocks (10 blocks for each of the emotion classes **VA**, **Va**, **vA**) of 35 seconds length. Each block contained 4-5 images and 4-8 sound samples. The order of the blocks and the selection of pictures and sounds was pseudorandomized. A neutral blank screen was displayed and no sounds were played for 30-40 seconds after each block and at the beginning of the experiment (called **Neutral** throughout the paper). The duration of the **Neutral** phases was randomly varied between 30 and 40 seconds to reduce a potential systematic influence of slow waves. The total recording time of one recording session was about 35 minutes. Figure 1 summarizes the experiment procedure.

D. fNIRS Measurement

To measure cerebral hemodynamics we used an Oxymon Mk III system by Artinis Medical Systems. The montage consists of four transmitter and four detector optodes. The optodes were attached to the subjects' forehead using a headgear, so that its lower edge was shortly above the eyebrows und the optodes were symmetrical to the head midline. Figure 2 illustrates the optode montage. Using this setup, the system outputs concentration changes in oxygenated (HbO_2) and deoxygenated (HbR) hemoglobin at 8 source-detector pairs (16 channels) at a distance of 3.5 cm using a sampling rate of 25 Hz. A light intensity sensor has been attached to the screen and recorded synchronously with the fNIRS data to assess exact timings of the pictures during the experiment.

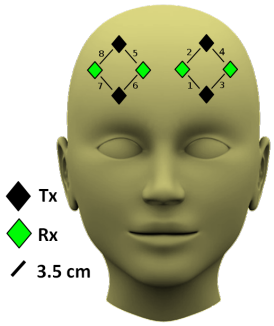

Fig. 2. Optode montage used in the experiment Tx stands for transmitters, Rx for detector optodes. The optode distances is 3.5 cm for all measurement locations.

III. METHODS

A. Signal Processing and Artifact Removal

The measured fNIRS signals are usually subject to biological and technical artifacts. Mostly influences from cardiovascular activity, such as heart beats and slow waves (e.g. Mayer Waves) as well as spikes caused by optode movements are present in the recorded signals (see e.g. [35], [36] for more detailed information on artifacts in fNIRS). To remove trends and slow waves, we applied a moving average filter, averaging over 180 seconds before and after each HbO_2 and HbR sample, and subtracted its output from the original signal. We applied an elliptic IIR lowpass filter with cutoff frequency 0.5 Hz (filter order 6) to remove higher frequency influences, including heart beat. Movement artifacts and spikes were removed using the wavelet based method proposed by Molavi and Dumont [37].

B. Feature Extraction

Several different feature extraction methods have been used for fNIRS based BCIs, including simple statistical properties of the time-domain signals, such as mean, standard deviation, slope, kurtosis and skewness. These features intuitively express properties of the characteristic shape of hemodynamic responses, when applied to stimulus-locked fNIRS signals. However, in the case of continuous recognition, most of these features do not reflect easily interpretable signal properties. We decided to use only the mean value of short windows of the HbO_2 and HbR signals for the evaluations in this paper, as preliminary experiments with other time-domain features did not show strong performance improvements. More advanced features, e.g. based on time-frequency transformations, might represent additional information of the signals, however, the mean has the advantages that it can very easily be calculated and interpreted. In previous studies [30], [32], the mean was also among the most successful feature to discriminate fNIRS signals in emotion tasks.

In order to continuously classify affective states, windows of 5 seconds length with 50 % overlap were extracted from the preprocessed fNIRS signals. Window lengths between 2 and 15 seconds were evaluated and appeared to give similar results, longer windows showed more unstable results, which can be explained by the block length of 35 seconds and small amount of data available. Each window was associated with

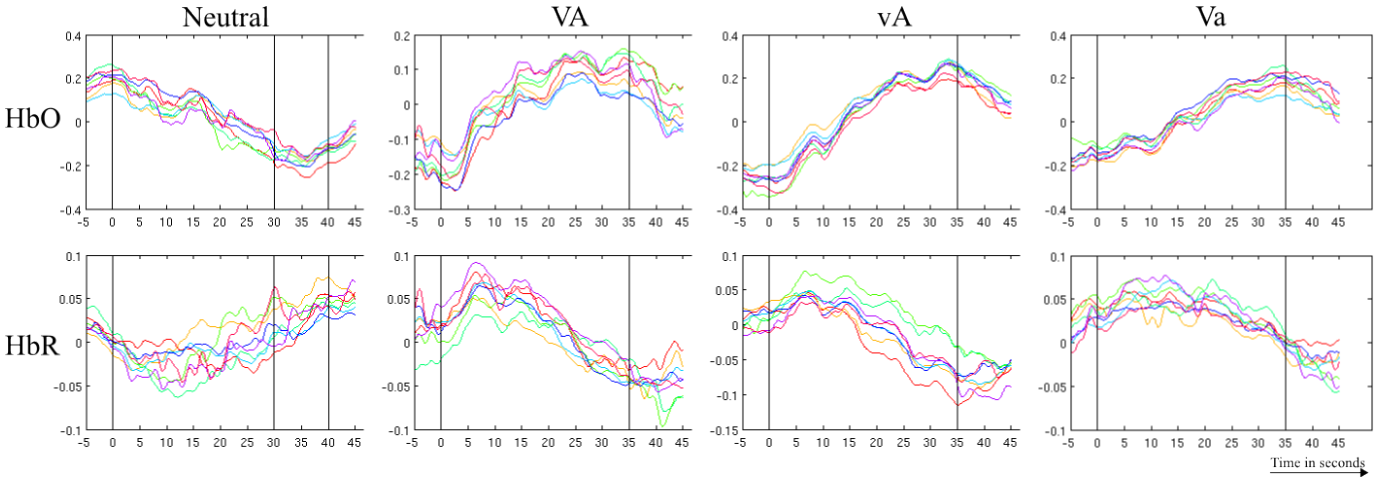

Fig. 3. Hemodynamic responses (grand averages) for the classes **VA**,**vA**,**Va**, and **Neutral**. The top row shows HbO_2 channels, the bottom row shows HbR channels. Signal epochs start 5 seconds before the beginning of an emotion induction or **Neutral** block and end 10 seconds after the end of the blocks. Vertical lines indicate begin and end of a block (as **Neutral** blocks have random length, the second and third vertical line in these plots indicate the time period where the blocks end). The x-axes show time in seconds, the y-axes concentration changes of HbO_2 and HbR.

its best fitting class label (**Neutral**, **VA**, **Va**, or **vA**).
We filtered out features that contained only little information using a mutual information based feature selection. For this purpose, mutual information between the continuous training data of each feature and the corresponding discrete emotion class labels (ground truth) has been calculated. Non-parametric probability density functions were estimated by kernel density estimation (Parzen windows). Features that contributed less than 10% to the total mutual information were excluded from the 16 dimensional feature vector.

C. Continuous Affect Classification

As the amount of data available to train and test the system is very limited, we decided to use 10-fold cross-validations for the evaluation of the classification performance of our system. Using this evaluation scheme, one has to be careful that no hidden information from the test data is used for training, otherwise the performance will be overestimated. The high temporal auto-correlation of fNIRS signals makes it problematic to include examples very shortly before or after the examples to be predicted in training, as neighboring training examples contain nearly the same information (e.g. this makes common evaluation schemes including random shuffling inappropriate).
In our cross-validation we split the data chronologically and ensured that there is a time gap of at least one minute between training and testing examples. In each fold of the cross-validation the training data was balanced to an equal number of examples for both classes, in order to avoid a bias of the classifier towards one class. Balancing was done by removing random examples from the training data. The mutual information based feature selection was performed on the training data in each fold of the cross-validation. We used support vector machines [38] with radial basis function kernels to classify the data. The penalty weight C for the slack variables and the kernel width γ were estimated using 10-fold cross-validation on the training data in each fold of the cross-validation ($C \in \{2^{-5}, 2^{-3}, \ldots, 2^{10}\}$, $\gamma \in \{2^{-15}, 2^{-13}, \ldots, 2^{3}\}$).

IV. Results and Discussion

A. Hemodynamic Responses to the Emotion Induction

In order to analyze the hemodynamic responses to the four different classes during emotion elicitation, we extracted stimulus-locked epochs starting 5 seconds before the beginning of an emotion induction block until 10 seconds after each emotion block. Figure 3 shows grand averages of the hemodynamic responses after applying the preprocessing as described in section III-A. Each plot in the upper row shows HbO_2 channels, the plots in the lower row show HbR channels averaged across subjects. The columns correspond to the classes **VA**, **vA**, **Va**, and **Neutral**, respectively. Hemodynamic responses to the emotion induction are obviously present and show a typical shape at all channels and all classes. For the three emotion classes (**VA**, **vA** and **Va**), HbO_2 channels show an increase (hyper-oxygenation), while HbR show a decrease in concentration few seconds delayed after the beginning of each block. One can see that the hemodynamic responses remain stable for the complete time of the emotion induction (block length 35 seconds) and turn towards baseline after the end of an emotion induction block. A contrary effect can be observed for the **Neutral** plots, i.e. HbO_2 decrease and HbR increase after the beginning of the block. This can be explained as **Neutral** blocks occur directly after the emotion stimulus blocks and include the decay of the hemodynamic activity.

B. Continuous Affect Recognition Performance

Figure 4 shows mean classification accuracies for all subjects and all binary combinations of the four classes (**Neutral**, **VA**, **vA**, **Va**) calculated using 10-fold cross-validation with a 1 minute gap between training and test sets. The error bars indicate standard deviations across the ten folds of the cross-validation. We calculated one-sided Wilcoxon signed rank tests to test if the recognition accuracies are significantly ($p < 0.05$) above the 50% chance level. All emotion classes (**VA**, **vA**, **Va**) could significantly be discriminated from **Neutral** for all subjects. The average performances across subjects were 65.4% (sd=4.0) for **Neutral** versus **VA**, 63.1% (sd=4.1) for

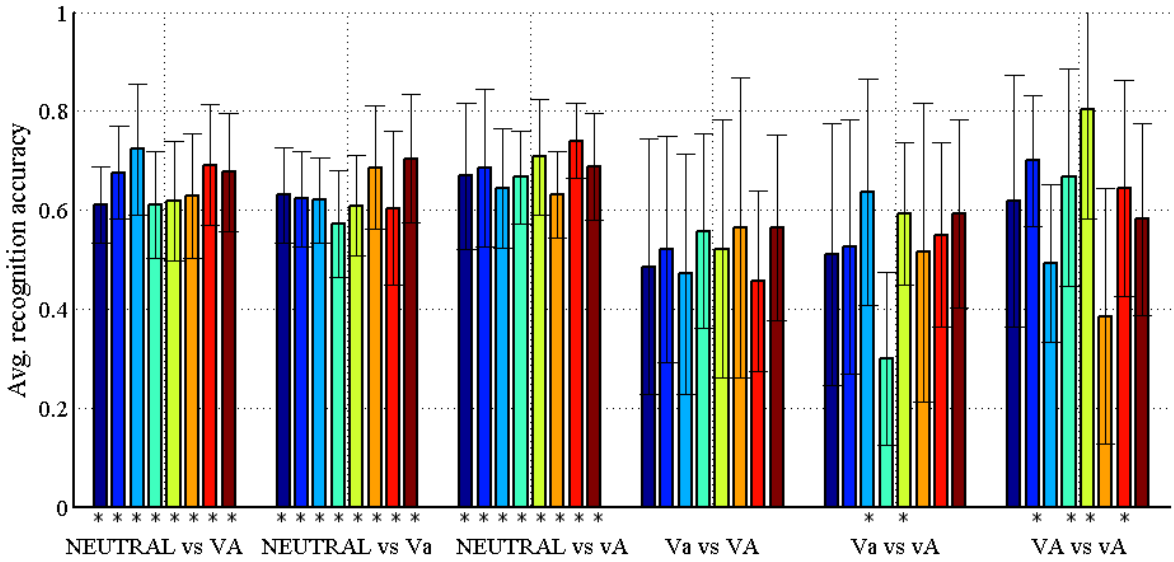

Fig. 4. Classification accuracies for all subjects and all binary classification tasks. The eight bars in each group correspond to the eight subjects. Error bars show standard deviation of five-fold cross-validation. Stars indicate results significantly above chance level of 0.5 ($p < 0.05$).

Neutral versus **Va**, and 67.9% (sd=3.2) for **Neutral** versus **vA**. Influences of the previous emotion block on the **Neutral** class were analyzed and none of the Wilcoxon rank sum tests showed significant differences between the classification accuracies of **Neutral**, depending on the class of the previous emotion block. The recognition rates in discriminating **VA** against **Va** were not significantly above chance level. This might be due to the fact that these classes are most similar according to the valence and arouse ratings in Table I. Subjects 3 and 5 achieved significant results in discriminating **Va** from **vA**. For **VA** versus **vA**, half of our subjects showed recognition rates significantly above chance level. This indicates that differences during high arousal result in more distinct activations than response involving low arousal. The average recognition rates across subjects were 51.8% (sd=4.0) for **Va** versus **VA**, 52.8% (sd=9.6) for **Va** versus **vA**, and 61.2% (sd=12.0) for **VA** versus **vA**.

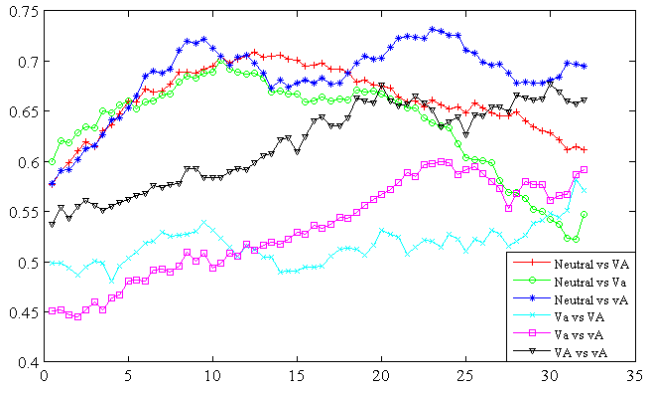

Fig. 5. Classification accuracies as a function of the time distance from the start of an emotion stimulation block. X-axis shows time offset in seconds, y-axis shows average recognition accuracies.

To analyze the performance of the continuous recognition in more detail, we calculated average recognition rates as a function of the time distance from the start of the corresponding emotion induction block. For this analysis the predictions from the cross-validation experiments (previous paragraph) were used. The time of each emotion induction and **Neutral** block was split into analysis intervals of 5 seconds length overlapping 4.5 seconds. The recognition accuracies of the predictions that occurred within each analysis interval were calculated (average across subjects). Figure 5 shows the resulting curves of all binary classification tasks. The x-axis shows the start time of the analysis intervals. All binary classification tasks show an increase in recognition accuracies after the begin of emotion induction. Surprisingly, recognition rates show a down trend for **Neutral** versus **VA** and **Neutral** versus **Va** starting at 10-15 seconds, which appears not to be supported by the plots in Figure 3. In all of the other recognition tasks the performance increased and remained rather stable. This analysis shows that the recognition results were not caused by short peaks after the emotion triggering event, instead, the continuous affective state classification using small windows is effective for classifying longer affective state periods. The peak values of the curves indicate that an optimal analysis interval would have increased average recognition rates above 70% for the discrimination of emotion classes versus **Neutral** in stimulus-locked evaluations.

V. Conclusion

In this paper, we presented a system for the continuous (i.e. asynchronous) recognition of affective states from fNIRS signals in response to different classes of audio-visual affective stimulation. In comparison to the results reported by previous studies, the performance of our system may appear inferior. However, continuous recognition is a much more challenging task than stimulus-locked evaluations. In this first study we chose a simple setup that allows an easy analysis and interpretation of all processing steps involved. Significant recognition results indicate that it is possible to discriminate **Neutral** signals from signals recorded during emotion induction. Discriminating different emotion classes against each other appeared to be more difficult. However, the results show that strongly

different classes may also be discriminated using continuous recognition. The analysis of classification performance over time indicates that the continuous recognition of affective states of longer duration (such as mood) might be possible as recognition rates appear to remain stable for most of the evaluated binary classification tasks. However, this should be verified by further studies that, for example, include emotion induction blocks of different lengths. In future work, most importantly, the performance of the recognition system needs to improved, for example by using individual optimizations for each subject, optode montages with additional locations, and advanced feature extraction methods. Further analyzes should also investigate influences of perceptual properties of the emotional pictures on the system performance in detail.

REFERENCES

[1] R. W. Picard, *Affective computing*. MIT press, 2000.

[2] S. K. DMello, S. D. Craig, A. Witherspoon, B. Mcdaniel, and A. Graesser, "Automatic detection of learners affect from conversational cues," *User Modeling and User-Adapted Interaction*, vol. 18, no. 1-2, pp. 45–80, 2008.

[3] D. Heger, F. Putze, and T. Schultz, "An eeg adaptive information system for an empathic robot," *International Journal of Social Robotics*, vol. 3, no. 4, pp. 415–425, 2011.

[4] R. e. Kaliouby, R. Picard, and S. BARON-COHEN, "Affective computing and autism," *Annals of the New York Academy of Sciences*, vol. 1093, no. 1, pp. 228–248, 2006.

[5] J. E. LeDoux, "Emotion circuits in the brain," *Annual review of neuroscience*, vol. 23, no. 1, pp. 155–184, 2000.

[6] G. Bush, P. Luu, and M. I. Posner, "Cognitive and emotional influences in anterior cingulate cortex," *Trends in cognitive sciences*, vol. 4, no. 6, pp. 215–222, 2000.

[7] A. R. Damasio, T. J. Grabowski, A. Bechara, H. Damasio, L. L. Ponto, J. Parvizi, and R. D. Hichwa, "Subcortical and cortical brain activity during the feeling of self-generated emotions," *Nature neuroscience*, vol. 3, no. 10, pp. 1049–1056, 2000.

[8] A. Bechara, H. Damasio, and A. R. Damasio, "Emotion, decision making and the orbitofrontal cortex," *Cerebral cortex*, vol. 10, no. 3, pp. 295–307, 2000.

[9] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general," *Journal of Neural Engineering*, 2011.

[10] G. G. Molina, T. Tsoneva, and A. Nijholt, "Emotional brain-computer interfaces," in *Affective Computing and Intelligent Interaction and Workshops. 3rd International Conference on*. IEEE, 2009, pp. 1–9.

[11] A. Sassaroli and S. Fantini, "Comment on the modified beerlambert law for scattering media," *Physics in Medicine and Biology*, vol. 49, no. 14, p. N255, 2004.

[12] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion recognition: a review," in *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on*. IEEE, 2011, pp. 410–415.

[13] M. El Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: Features, classification schemes, and databases," *Pattern Recognition*, vol. 44, no. 3, pp. 572–587, 2011.

[14] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (iaps): Technical manual and affective ratings," 1999.

[15] M. M. Bradley and P. J. Lang, "The international affective digitized sounds (; iads-2): Affective ratings of sounds and instruction manual," *University of Florida, Gainesville, FL, Tech. Rep. B-3*, 2007.

[16] P. Ekman *et al.*, "Facial expression and emotion," *American Psychologist*, vol. 48, pp. 384–384, 1993.

[17] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of research in Personality*, 1977.

[18] P. Lang and M. M. Bradley, "The international affective picture system (iaps) in the study of emotion and attention," *Handbook of emotion elicitation and assessment*, p. 29, 2007.

[19] M. Herrmann, A. Ehlis, and A. Fallgatter, "Prefrontal activation through task requirements of emotional induction measured with nirs," *Biological psychology*, vol. 64, no. 3, pp. 255–263, 2003.

[20] Y. Hoshi, J. Huang, S. Kohri, Y. Iguchi, M. Naya, T. Okamoto, and S. Ono, "Recognition of human emotions from cerebral blood flow changes in the frontal region: A study with event-related near-infrared spectroscopy," *Journal of Neuroimaging*, vol. 21, no. 2, 2011.

[21] K. Morinaga, J. Akiyoshi, H. Matsushita, S. Ichioka, Y. Tanaka, J. Tsuru, and H. Hanada, "Anticipatory anxiety-induced changes in human lateral prefrontal cortex activity," *Biological Psychology*, vol. 74, no. 1, pp. 34 – 38, 2007.

[22] J. Leon-Carrion, J. Damas, K. Izzetoglu, K. Pourrezai, J. F. Martín-Rodríguez, J. M. B. y. Martin, and M. R. Dominguez-Morales, "Differential time course and intensity of pfc activation for men and women in response to emotional stimuli: a functional near-infrared spectroscopy (fnirs) study," *Neuroscience letters*, vol. 403, no. 1, pp. 90–95, 2006.

[23] H. Yang, Z. Zhou, Y. Liu, Z. Ruan, H. Gong, Q. Luo, and Z. Lu, "Gender difference in hemodynamic responses of prefrontal area to emotional stress by near-infrared spectroscopy," *Behavioural brain research*, vol. 178, no. 1, pp. 172–176, 2007.

[24] E. Glotzbach, A. Mühlberger, K. Gschwendtner, A. J. Fallgatter, P. Pauli, and M. J. Herrmann, "Prefrontal brain activation during emotional processing: A functional near infrared spectroscopy study (fnirs)," *The open neuroimaging journal*, vol. 5, p. 33, 2011.

[25] A. Ishikawa, K. Shimizu, and N. Birbaumer, "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface," *NeuroImage*, 2007.

[26] S. M. Coyle, T. E. Ward, and C. M. Markham, "Brain–computer interface using a simplified functional near-infrared spectroscopy system," *Journal of neural engineering*, vol. 4, no. 3, p. 219, 2007.

[27] A. Girouard, "Adaptive brain-computer interface," in *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2009.

[28] K. K. Ang, C. Guan, K. Lee, J. Q. Lee, S. Nioka, and B. Chance, "A brain-computer interface for mental arithmetic task from single-trial near-infrared spectroscopy brain signals," in *Pattern Recognition (ICPR), 20th International Conference on*. IEEE, 2010, pp. 3764–3767.

[29] C. Herff, F. Putze, D. Heger, C. Guan, and T. Schultz, "Speaking mode recognition from functional near infrared spectroscopy," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 1715–1718.

[30] K. Tai and T. Chau, "Single-trial classification of nirs signals during emotional induction tasks: towards a corporeal machine interface," *Journal of neuroengineering and rehabilitation*, vol. 6, p. 39, 2009.

[31] S. Hosseini, Y. Mano, M. Rostami, M. Takahashi, M. Sugiura, and R. Kawashima, "Decoding what one likes or dislikes from single-trial fnirs measurements," *Neuroreport*, vol. 22, no. 6, p. 269, 2011.

[32] S. Moghimi, A. Kushki, S. Power, A. M. Guerguerian, and T. Chau, "Automatic detection of a prefrontal cortical response to emotionally rated music using multi-channel near-infrared spectroscopy," *Journal of Neural Engineering*, vol. 9, no. 2, p. 026022, 2012.

[33] H. Asano, T. Sagami, and H. Ide, "The evaluation of the emotion by near-infrared spectroscopy," *Artificial Life and Robotics*, pp. 1–5, 2013.

[34] C. Amma, A. Fischer, T. Stein, H. Schwameder, and T. Schultz, "Emotionserkennung auf der basis von gangmustern," in *Sportinformatik trifft Sporttechnologie, Tagung der dvs-Sektion Sportinformatik*, 2010.

[35] F. Matthews, B. Pearlmutter, T. Ward, C. Soraghan, and C. Markham, "Hemodynamics for brain-computer interfaces," *Signal Processing Magazine, IEEE*, vol. 25, no. 1, pp. 87–94, 2008.

[36] R. Cooper, J. Selb, L. Gagnon, D. Phillip, H. Schytz, H. Iversen, M. Ashina, and D. Boas, "A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy," *Frontiers in Neuroscience*, vol. 6, p. 147, 2012.

[37] B. Molavi and G. A. Dumont, "Wavelet-based motion artifact removal for functional near-infrared spectroscopy," *Physiological measurement*, vol. 33, no. 2, p. 259, 2012.

[38] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.