# A data-driven validation of frontal EEG asymmetry using a consumer device

Doron Friedman
The Advanced Reality Lab
The Interdisciplinary Center
Herzliya, Israel
doronf@idc.ac.il

Shai shapira, Liron Jacobson, Michal Gruberger
Brainster, Inc.
Kfar Saba, Israel

***Abstract*— Affective computing requires a reliable method to obtain real time information regarding affective state, and one of the promising avenues is via electroencephalography (EEG). We have performed a study intended to test whether a low cost EEG device targeted at consumers can be used to measure extreme emotional valence. One of the most studied frameworks related to the way affect is reflected in EEG is based on frontal hemispheric asymmetry. Our results indicate that a simple replication of the methods derived from this hypothesis might not be sufficient. However, using a data-driven approach based on feature engineering and machine learning, we describe a method that can reliably measure valence with the EPOC device. We discuss our study in the context of the theoretical and empirical background for frontal asymmetry.**

***Keywords*— *EEG, emotion; valence; frontal asymmetry***

## I. Introduction

Affective computing [1] suggests a whole range of exciting applications, but obtaining reliable probes into affective state remains a challenge. Similar paradigms are now studied under fields such as physiological computing (Fairclough, 2009) and passive brain computer interfaces (Zander and Kothe). Electroencephalography (EEG) provides a promising avenue for extracting affective state information directly from the brain. In the last few years several low cost EEG devices were made available that are intended for use outside the laboratory. If such devices can be used for affective computing applications this could facilitate a breakthrough in the adoption of such applications outside research laboratories.

Recently, there have been some efforts at automatic recognition of emotions from EEG, but this has proved challenging. The most basic theoretical model of affect is based on valence and arousal ([2], [3]). Levels of arousal and valence can be extracted from autonomous nervous system (ANS) recordings such as electrodermal activity or heart rate, but recognizing emotions from ANS is a major challenge, and a matter of ongoing debate in psychophysiology [4]. It is reasonable to expect that the information extracted from the central nervous system would be richer, and it is now clear that in humans the cortex plays a large role in emotional processing [5], making this type of information available to EEG, at least in principle. However, whether this is also the case in practice is still a matter for debate in the neurophysiology community, e.g.,: "perhaps cortical EEG patterns will never be able to be used to distinguish discrete emotional states from the surface of the brain" [6](p. 1959).

We have developed a method for online measurement of emotional information using a consumer EEG device (Emotiv EPOC[1]). We have carried a study aimed at reliably distinguishing positive from negative valence, while subjects were watching video clips. Our findings demonstrate that despite the high percentage of noise in the EPOC EEG data it is possible to recognize temporal segments with extreme valence using a data-driven approach with reasonable accuracy, and the results are consistent with the neurophysiological literature.

## II. Background

### A. *EEG correlates of affect*

A meta review [7] of functional magnetic resonance imaging (fMRI) concludes that there are clear anatomical signatures for at least five basic emotions: happiness, sadness, anger, fear, and disgust. It is clear that emotion is processed in subcortical areas, but there is an ongoing debate to what extent the cortex is involved in this processing [8], [9].

There has been many studies related to affective state and EEG, but the results are often inconsistent and scattered. An exception is the line of work suggesting that emotional valence is related to hemispheric asymmetry, mostly in the prefrontal cortex. First evidence came from lesion and brain trauma studies: unilateral lesion indicated that left hemisphere trauma results in depressive reaction whereas right hemisphere trauma results in euphoric reaction [10]. Robinson et al. [11] continued to suggest that depression is a result of left frontal trauma, and posterior region trauma does not have the same result. In the right hemisphere posterior damage resulted in depression, whereas in the case of frontal damage patients were more cheerful. This line of research continued with studies with healthy subjects, showing greater right hemisphere involvement in emotional processing [12]. In a series of studies Davidson and colleagues asked people to provide like/dislike ratings while watching TV clips; liked segments resulted in an increase in right frontal alpha, and disliked in left increase in alpha (e.g., [13]).

There are four models related to frontal asymmetry in brain activations, with slightly different emphasis and predictions [Demaree 2005]. The 'right hemisphere model' suggests that emotion perception and expression are largely subserved by the right cerebrum. Two models highlight left versus right activation: the valence model and the approach-withdrawal. A fourth model is the behavioral-inhibition system behavioral

[1] http://www.emotiv.com

activation system (BIS-BAS) model. The latter is focused on traits rather than states. While inter-individual differences are clearly significant and important, in many affective computing applications we cannot except to take such differences into accounts. Thus, below we will focus on the valence and the approach-withdrawal models.

The alpha rhythm is considered to be correlated with a decrease in neural processing, so a simplified statement of the valence hypothesis is that positive emotions are associated with increased processing in the frontal left hemisphere and negative emotions are associated with increased frontal right hemisphere processing. A simplified description of the approach-withdrawal hypothesis is that increased left frontal activity is associated with approach related behaviors and increased right frontal activity is associated with withdrawal related behaviors. Anger, despite being a negative valence emotion, is considered an approach related behavior. Since anger has been found to correlate with increased left frontal activity, as well as for other reasons, the approach-withdrawal hypothesis is considered to have more support than the valence hypothesis [5], [13], [14].

The prefrontal cortex (PFC) is diverse and complex. Using animal studies and functional magnetic resonance imaging (fMRI) there are now large amounts of high resolution data and elaborate theoretic models. Our challenge is that EEG is limited: it cannot detect subcortical regions (although see [15] for a promising approach based on synchronous fMRI-EEG recordings), and it has poor spatial resolution. Based on an increasing understanding of cortical and subcortical processing of emotions it is possible to predict the type of information that may be accessible by scalp EEG recordings. For example, Davidson et al. predict that emotional tasks that include a response component are more likely to be reflected in EEG [16]; this is good news for interactive applications. They also suggest that the orbital PFC `computes' affect, but this is not accessible to EEG [16]. Harmon-Jones et al. [17] report that left frontal activation only occurred when coping or retaliatory responses were possible; again this is good news for interactive applications, as opposed to passive viewing such as in our current study.

Frontal asymmetry is both an individual difference variable and a state variable. Coan et al. [18] highlight the distinction between moderating and mediating variables. If frontal asymmetry only acts as a moderator, i.e., it determines what response an individual would have to a stimulus – it is not appropriate as an affective index. Frontal EEG asymmetry is said to mediate emotional responses to the degree that is serves as, or is highly correlated with, the mechanism by which emotional stimuli exert their effects. There are several studies suggesting that frontal asymmetry also acts as a mediator [19]–[21]. Specifically, Coan et al. [19] found a g-coefficient (in a Cronbach g-analysis) of 0.97 for state frontal asymmetry – to the extent that state changes in frontal EEG asymmetry occurred at all in response to the emotional manipulation, they occurred in nearly all of the subjects, regardless of their trait predispositions.

There is also some evidence that the frontal theta band plays a role in affective processing, regardless of the asymmetry framework. Bekkedal et al. [6] detected changes in theta, and specifically frontal theta, during early emotional processing. They suggest that cortical theta is an extension of the theta activity in limbic systems, mostly hippocampus and amygdala. Vecchaito et al. [22] also report a significant increase in frontal theta for pleasant commercials. In another study [23] the same group also reports alpha asymmetry in consistence with Davidson's asymmetry framework.

A major challenge in uncovering the neural correlates of emotion is the methodological challenge of eliciting emotions. Most scientists call for 'clean' minimalistic stimuli, in order to disentangle emotional factors from other factors; for example Bekkedal et al. suggest using only auditory stimuli rather than multimodal stimuli such as movies [6]. However, their findings as well as others highlight the methodological tradeoff: their experimental manipulation did not result in a subjective reporting of emotions. In the context of most affective computing applications the decision is easy – since eventually we want the metrics to be used in rich media multimodal context, we should learn from experiments with minimalistic stimuli, but we should also use rich stimuli such as movie clips.

### B. *Machine learning approaches*

More recently there are growing efforts to use a data driven, machine learning approach for extracting affective state from EEG, and our approach falls in this basket. Nevertheless, we suggest that it is not only of interest but also necessary to explain the machine learning results in terms of the underlying neurophysiological mechanisms.

Jenke et al. [24] provide an extensive review of features as well as feature extraction methods for inferring emotion from EEG. They collected methods from 33 studies and compared them on their dataset of 16 subjects (11 after removing bad recordings). The dataset included 8 trials of 30s EEG recordings for five emotions: happy, curious, angry, sad, and quiet. Emotions were induced using the international affective picture system (IAPS) image collection. Subjects self-reported their emotions while watching 160 emotional pictures covering the arousal valence dominance space, using the self-assessment manikin (SAM) questionnaire [2]. EEG was recorded from 64 channels. They report relatively low accuracy rates (25-47.5% relative to 20% chance) and high variability among subjects, features, and methods. They conclude that complex features such as high order crossing (HOC) or features measuring the signal complexity in the time domain are most valuable, whereas spectral power gave low results. In addition, they suggest that multivariate feature selection methods (such as maximum relevance minimum redundancy, mRMR) are superior to univariate methods (such as effect size, $f^2$) for feature selection.

Khosrowabadi et al. [25] recognize four emotions: happiness, sadness, fear, and calm, with 84.5% accuracy. They also used IAPS but enhanced the experience with music to induce the emotions. They used features extracted from magnitude squared coherence estimates (MCSE); i.e., the correlation between electrodes in the frequency domain. They show that a significant improvement in results was achieved by using self-organizing maps to cluster the subjective reports (questionnaires) into four groups corresponding to the emotions.

Petrantonakis et al. [26] achieved up to 83.3% distinguishing six basic emotions. Their results are based on a small set of frontal EEG channels (FP1, FP2, and bipolar F3-F4). Valenzi et al. [27] report up to 97.2% classification accuracy, with very little inter-subject variance, in distinguishing four emotions: amused, disgusted, sad, and neutral. They use 16 video clips between 40s and 324s from famous movies. The clips were subdivided into epochs of 6 seconds, and band power was used as features. Using linear discriminant analysis (LDA) they reduced 160 features for each epoch into 3. Interestingly, their results support the frontal asymmetry hypothesis.

Koelstra et al. [28] selected 40 music video clips ranked with extreme values of arousal and valence, measured 32 subjects' EEG and ANS activity, and made the dataset publicly available. They tested Spearman correlation between the power of each of the 32 electrodes in the main frequency bands and subjective ratings, and report that positive valence was correlated with an increase in alpha and theta in occipital areas, but indicate that this could be de-activation due to focus on sound. The correlations are very low - seldom larger than +-0.1. As a next step, they attempted single trial classification of low/high arousal, low/high valence, and low/high liking with the same dataset, using subject ratings as ground truth. The features are the power in frequency bands and also the difference between symmetric electrodes (this is similar to our approach). They report moderately successful results – EEG classification results are significantly better than chance but are less than 60% (F-scores) for arousal and valence, and no classification was achieved for liking.

We suggest that the large variation of success among these studies, ranging from no classification to 97% classification, may be attributed to the psychological factors of the experimental design, such as the type of stimuli used for emotional induction.

### C. Consumer EEG devices

The studies reported above can be considered encouraging, but they have all been obtained in laboratory experiments, using professional EEG devices that suffer from a prohibitive cost and reliance on expertise. Consumer EEG devices, which are low cost and easy to use, can lead to a revolution in affective computing, and in human computer interface (HCI) in general. However, devices such as the Emotiv EPOC generate data that is highly contaminated with various sources of noise. Some of the artifacts, such as eye blinks and facial muscle activity can be very useful, and in fact Emotiv supplies this information as an 'affective suite', and this may be a promising avenue for affective computing. In this paper, however, we are only interested in information extracted directly from the brain via EEG.

Studies with the EPOC seem to agree that it can be used to acquire reasonable EEG signals, but it is of lower quality (e.g., [29][30]), and our conclusions are similar. Khushaba et al. [31], [32] performed two studies on user preferences with the EPOC device. They do not explicitly study emotional valence but their results are very encouraging in using the EPOC for affective computing.

There have been some other preliminary studies using the EPOC for affective computing. Liu et al. [33] suggest an affective computing application, whereby an advertisement clip automatically changes based on levels of valence, arousal, and dominance extracted from the EEG EPOC. Such adaptive content is indeed the direction our research aims at; however, Liu et al. do not provide any experimental results. Ramirez et al. [34] report a study that is very similar to ours, and their results are also in line with ours; they detect arousal and valence with accuracies of approximately 80%. Our paper goes beyond this study in several directions: we use a much larger subject pool (29 subjects, whereas their study is based on 6 subjects), and we use a more data-driven approach, including an investigation of feature selection, intended to probe the frontal asymmetry hypothesis. Additionally, in their study Ramirez et al. address the issue of noise in the EPOC signals by assuming that the alpha and beta bands are not contaminated – our experience indicate that this is not realistic; the abundance of artifacts results in various types of unexpected shifts to the frequency spectrum.

## III. Method

### A. Subjects

The study included 44 subjects (25 females, mean age = 36), recruited in a local shopping mall, and paid the equivalent of $15 for participation.

### B. Setup

The subjects were seated on a chair in a distance of approximately 70cm in front of a 25 inch monitor. A technician fitted the EPOC system on their scalp. The EPOC device does not include standard electrode locations, and thus accurate reconstruction of experiments is difficult; the electrodes roughly match the following international 10-20 standard positions: AF3, AF4, F3, F4, F7, F8, FC5, FC6, T7, T8, P3, P4, P7, P8, O1, and O2. Two additional electrodes are used as reference electrodes and their location can be slightly modified, either below or above the ears. We recorded above the ears since recording below the ears often resulted in pulse artifacts.

### C. Procedure

Subjects were given an explanation regarding the experimental procedure, and were asked to fill in a consent form and a demographics questionnaire. The study included a sequence of video clips with baseline intervals of 10 seconds between the clips. During the baseline the screen was gray and the subjects were asked to look at a swirling 'x'; this was selected as an emotionally neutral visual stimulus.

The study included 12 video clips with durations of approximately one minute. Most of the clips were local commercials, and three out of the clips were extremely emotional short clips; two were extremely negative, one was extremely positive. The reason for using commercials was that this study was sponsored by a company that is currently interested in neuromarketing (see, e.g., [35], for background on neuromarketing). The clips appeared in a different and arbitrary order for the different subjects.

### D. Data Analysis

#### 1. Subjective Selection of Extreme Segments

Ground truth is a major challenge for measuring emotional and cognitive state [16]. Eight separate subjects, referred to as coders, with demographics similar to that of the experiment group, were asked to watch all video clips and rate the most extreme segments in terms of valence, all having high arousal. The instructions to the subjects were to mark the timing of peak moments – short segments that made them feel either extremely positive or extremely negative emotions. In total these subjects marked 95 positive valence segments and 66 negative valence segments, with a mean duration of 5.76 seconds (std = 5.54). Out of these we have selected 17 segments based on a criteria of 75% inter-coding agreement or higher: 10 segments rated positive and 7 segments rated negative. Thus, by definition there is a significant difference between the two categories of clips in terms subjective reporting. Further studies are needed to refine the type of emotions and to see whether a continuous ranking can be correlated with EEG.

#### 2. Preprocessing and Artifact Rejection

The data was sampled by the EPOC device; according to the manufacturer the sampling is done internally at 2048 Hz and down sampled to 128 Hz. We used a high-pass Butterworth filter of order 4 with a cutoff frequency of 2 Hz (low frequencies are especially contaminated with noise) and a low-pass Butterworth filter of order 4 with a cutoff frequency of 43 Hz.

The EEG data was manually inspected and noise and artifacts were marked and removed by an EEG expert. This included removing whole channels for some recordings and short temporal segments over all channels. If the baseline period before a specific clip was contaminated then the data of that subject from that clip was unusable, since the baseline is required for normalization. Overall 15 out of 44 subjects were completely removed from the study due to very noisy recordings (many disconnected electrodes), and from the rest we have removed 41% of the data samples. In the analysis of extreme segments we further removed samples with more than 25% noise. The percentage of noise is quantified based on the number of samples removed, out of overall $CxT$ samples, where $C$ is the number of channels (14) and $T$ is the number of samples per clip (128 Hz multiplied by number of seconds in the clip). This percentage of noise is unacceptable in academic experiments, but might be coped with in interface applications in the real world.

#### 3. Feature Extraction

All our analysis is based on extracting frequency domain features from the raw EEG signals. The first set of features were the values of the power in the common EEG frequency bands, extracted with a fast Fourier transform (FFT): delta (1-4 Hz), theta (4-7 Hz), alpha (8-13 Hz), low beta (13-21 Hz), high beta (22-32 Hz), and root mean square (RMS, 2-40 Hz). These features were computed for the mean of the whole clip duration, as well as for the duration of the extreme segments, and were normalized by extracting the value of the features in the baseline period before each clip. We did not add the gamma band because it is not expected to be reliable in consumer devices such as EPOC, due to muscle artifacts. Also, we do not report results in the delta band since it seems to be most contaminated with noise, such as eye movement artifacts.

A second set of features was extracted for hemispheric differences: the differences in band power in the main frequencies, between all pairs of symmetric electrodes, left hemisphere – right hemisphere (e.g., AF3-AF4, T7-T8). In addition, this set of features included the difference in band powers between frontal electrodes (AF3+F3+F7+FC5) – (AF4+F4+F8+FC6), the difference in band powers between the posterior electrodes over the two hemispheres (T7+P7+O1) – (T8+P8+O2) and the difference in band powers in the 7 left hemisphere electrodes and the 7 right hemispheres electrodes.

## IV. Results

### A. Statistical Analysis

We have performed a multivariate analysis of variance (MANOVA) with Bonferonni correction for multiple comparisons, with independent variables SUBJECT and SEGMENT, and the multiple features described in Section III.d.iii as dependent variables. In the first test the features were the mean values from four clips: two rated highest and two rated lowest in valence, by all subjects. This analysis indicated that there were no significant differences between the two valence categories.

In the second step we repeated the analysis but this time with the means of the extreme segments revealed by the coding process. Figure 1 provides a schematic view of the results taking into account the power in the different frequency bands in single electrodes. Significant differences were found mostly in the alpha range, and roughly corresponds with the asymmetry hypothesis; unlike what would have been expected from the framework we have found evidence in non-frontal electrodes, and an inverse effect was found for occipital electrode O1 in the alpha range.

The MANOVA analysis can also be used as a dimensionality reduction method, since it finds a linear combination of the original variables (referred to as features) that yields the largest separation between groups (classes). Figure 2 provides a visualization of the top two variables resulting from the MANOVA analysis, indicating that it provides a reasonable separation between the classes.

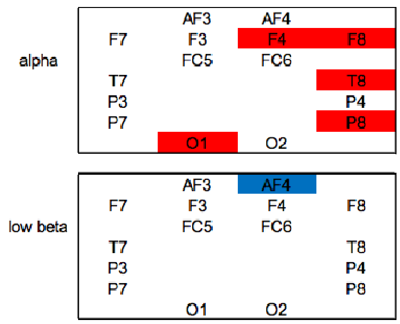



Figure 1: A view of the MANOVA results over all frequency bands and all electrodes, comparing segments with extreme reported valence. The locations of the EPOC device cannot be placed with high accuracy, so we present a schematic view rather than scalp topography maps. An electrode is highlighted in red if the power in the corresponding band was significantly higher in positive valence clips as opposed to negative valence clips. An electrode is highlighted in blue if the power in the corresponding band was significantly lower in positive valence clips as opposed to negative valence clips. Electrodes that are not highlighted did not distinguish among the two valence categories. No significant founding was found in this analysis in other frequency bands.

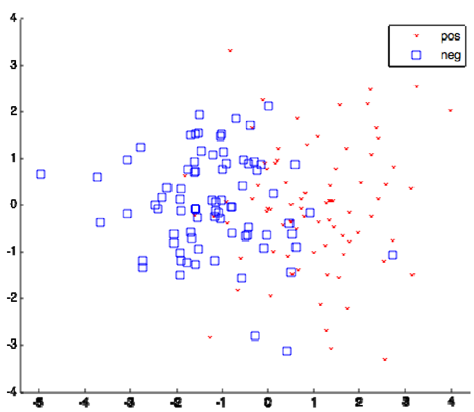



Figure 2: A visualization of the first two variables of the MANOVA analysis, indicating the reasonable separation between high and low valence. 'Pos' denotes positive valence clips and 'neg' denotes the negative valence clips.

### B. *Single Sample Classification*

For affective computing it is not enough to uncover correlations between affect and neural signals; we must be able to measure the level of affect in a single trial, or to recognize points in time with extreme affect online.

The large number of features and the small number of samples results in a typical case of 'over fitting'. Our study resulted in 229 labelled samples – 125 positive and 104 negative, with 163 features. Therefore, we compare several means of feature selection: using all features, selecting only features that were found significant in the MANOVA test above (Figure 2), and using the information gain algorithm [36] for feature ranking and selection. Since the MANOVA test found only 6 features to be significantly different in both conditions, we have extended this set to include nearly significant results ($p < 0.08$), resulting in 30 features.

For classification we use the sequential minimal optimization (SMO) [37] implementation of the popular support vector machine (SVM) algorithm ($C = 1$, polynomial kernel), and we have also compared our results with other classification algorithms. In this case we have found that indeed SVM provides good results with our dataset but the logistic model trees (LMT) algorithm [38] provided slightly better classification rates. We have used the Weka implementation [39] of classification algorithms, and a combination of our own code and Weka for feature selection.

The results reported below are all based on nested (feature selection and classification) 10-fold cross validation applied on the whole dataset. The number of samples was not completely balanced between categories (55% positive and 45% negative samples), so we also report the F-Score values. The best classification result, 81.2% (F-score 0.812) was obtained with the LMT algorithm using all features suggested by the MANOVA. This cannot be achieved by chance; a t-test comparing the classification results versus a simulation (as suggested by Scherer and Brunner, 2008) indicates that 80% accuracy clearly falls outside the confidence bounds of a 95% certainty, which are around 60% accuracy in our case. The Kappa statistic (0.59) and MCC statistic (0.61) associated with this classification also indicate a strong positive correlation between the ground truth and the classifier results.

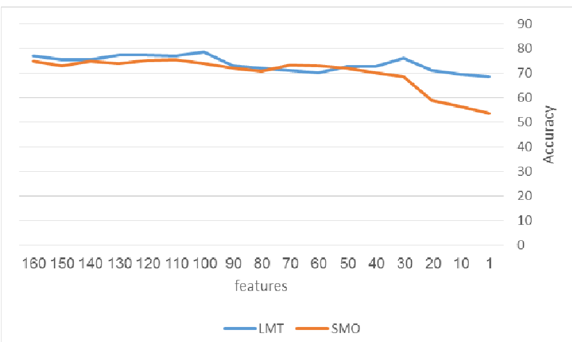
Table 1 provides a comparison of results using different techniques. While the differences are not dramatic, we see that using the features selected by MANOVA is usually optimal. When using LMT the combined set of features yielded the best result, while using SMO the single electrode features yielded the best result.

Using information gain for feature selection, our results indicate that almost all features had a positive information gain (above 0.2), and only a few of the features were discarded (information gain = 0). In addition, we see that the symmetric electrode differences are not very informative by themselves – this is evident by classification rates near chance (up to an F-score of 0.59), whereas the other type of features yielded above chance results (F-score values of 0.73-0.81).

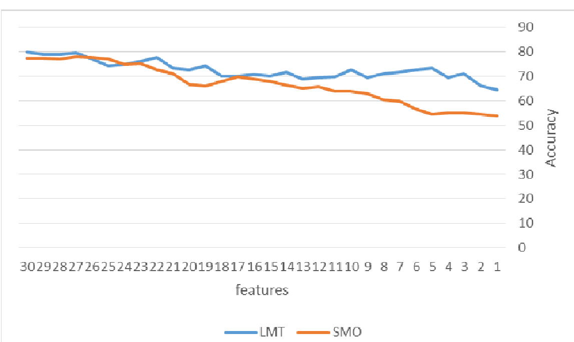
Figure 3 shows how the classification accuracy is affected by the number of features. Using SMO, it seems that adding features gradually contributes to the result. Using LMT classification we see that a single feature is enough for above chance classification (68%) – low beta in electrode AF3. Although this feature was not found significant in the MANOVA analysis (p=0.07), it has the highest information gain. Finally, we estimate the contribution of various types of features according to their classification rate (Table 2). These results seem to support the literature in revealing frontal electrodes as the main source of information, but as can be seen the full story is more complex.

TABLE I: CLASSIFICATION RESULTS COMPARING the SMO and LMT ALGORITHMS, THREE METHODS OF FEATURES SELECTION: USING ALL FEATURES (ALL), ONLY FEATURES WITH POSITIVE INFORMATION GAIN (INFOGAIN), AND ONLY FEATURES THAT WERE FOUND SIGNIFICANT USING A MANOVA TEST, AND THREE TYPES OF FEATURES: ONLY SINGLE ELECTRODES (SINGLE), ONLY THE POWER DIFFERENCE BETWEEN SYMMETRIC PAIRS OF ELECTRODES (SYM) AND THE COMBINED FEATURES (ALL).

| | | MANOVA | | IG | | all | |
|---|---|---|---|---|---|---|---|
| alg | | F-score | % | F-score | % | F-score | % |
| SMO | single | 0.77 | 77.3 | 0.74 | 74.7 | 0.73 | 73.4 |
| SMO | sym | 0.54 | 59 | 0.42 | 52.4 | 0.53 | 55.9 |
| SMO | all | 0.74 | 74.7 | 0.75 | 74.7 | 0.74 | 74.2 |
| LMT | single | 0.77 | 77.3 | 0.77 | 77.7 | 0.77 | 77.3 |
| LMT | sym | 0.55 | 57.2 | 0.53 | 56.3 | 0.59 | 61 |
| LMT | all | 0.81 | 81.2 | 0.78 | 78.2 | 0.75 | 74.7 |




(a)




(b)

Figure 3: Classification accuracy as a function of the number of features, comparing the SMO and LMT algorithms. The features are ranked by information gain. a) using all features, b) using only features selected by the MANOVA analysis.

## V. Discussion

Overall, our results indicate that hemispheric asymmetry may be used as a good marker for emotional valence in EEG, even when using a consumer device. The analysis is mostly based on extreme segments, for which high arousal was reported, so it is likely that our results indeed relate to the valence component of the emotional experience of our subjects. Our results are in line with the classic studies by Davidson et al. (e.g., [40]); on average there was no significant difference in frontal EEG when averaging results of whole clip duration, only when short segments were taken into account. Our results indicate that simple approximations of frontal asymmetry often used, such as $\alpha_{F4}$-$\alpha_{F3}$, or other types of inter-hemispheric differences, do not provide clear results in our dataset. Our results indeed indicate that frontal electrodes provide most information, but the findings only roughly fit the asymmetry model (e.g., consider the deviation in Figure 1, Table 2). Thus, a machine learning approach seems essential.

Importantly, our results were obtained with a low cost consumer device. This is not just a matter of cost. In order for end users to adopt brain-based applications, such as those in affective computing, devices need to be designed to be friendly and easy to use, and the EPOC makes important steps in this direction. Our conclusion is that although the levels of noise and artifacts are extremely high (Section III.D.2) it is nevertheless possible to extract useful information, in this case emotional valence, from such devices.

The drawback of the consumer device was mostly in its maintenance – it is not very durable, the electrodes suffer from corrosion, and the contact of the electrodes is problematic. Additionally, we note that the recordings were performed by lab technicians. This is an intermediary step towards real world deployment – the technicians are not EEG experts and were trained very briefly. Nevertheless, it is not clear if recordings carried out by complete novices would have any value at all.

The road towards EEG-based affective computing includes additional steps. In this study we show single trial classification, but only for extreme cases of emotional valence. Data was analyzed for multiple subjects, offline, whereas affective computing requires online single trial classification, which is more ambitious. A metric based on gradual ranking is required, or at least being able to classify a third, null class, of neutral valence. Continuous measurement over time (Section IV.c) requires further validation as well.

Can we recognize more subtle nuances of emotional experience? Bekkedal et al. [6] conclude their review paper as follows: (p. 1967): "Is it possible to ever get a clean EEG signal from surface recording that might highlight the neurodynamics of basic emotional processes in the brain? We suspect that

TABLE II: CLASSIFICATION RATES USING LMT COMPARING DIFFERENT SETS OF FEATURES.

| | % | F-score |
|---|---|---|
| Only α | 70.7 | 0.704 |
| F3 and F4 electrodes | 72.0 | 0.718 |
| No frontal electrodes | 75.5 | 0.755 |
| Only frontal electrodes | 79.0 | 0.794 |

may not be possible." Their pessimism is based on high inter-individual differences and on the idiosyncratic nature of emotional responses. Our suggestion is that fields such as affective computing can potentially overcome this challenge, by using machine learning, which can be tailored to take into account this high inter-subject variability.

## ACKNOWLEDGMENT

The study was partially funded by Brainster, Inc., a company developing neuromarketing solutions. No commercial interests affected the performance of this study. We wish to thank Isaac Blau, Noam Hacham and Maria Markovic for their support in EEG recording and data collection.

## REFERENCES

[1] R. W. Picard, *Affective Computing*, vol. 73. 1997.

[2] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm, "Looking at pictures: affective, facial, visceral, and behavioral reactions.," *Psychophysiology*, vol. 30, no. 3, pp. 261–273, 1993.

[3] J. A. Russell, "A circumplex model of affect.," *Journal of Personality and Social Psychology*, vol. 39, no. 6. pp. 1161–1178, 1980.

[4] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological Psychology*, vol. 84, no. 3. pp. 394–421, 2010.

[5] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion.," *Brain Cogn.*, vol. 20, no. 1, pp. 125–151, 1992.

[6] M. Y. V Bekkedal, J. Rossi, and J. Panksepp, "Human brain EEG indices of emotions: Delineating responses to affective vocalizations by measuring frontal theta event-related synchronization," *Neuroscience and Biobehavioral Reviews*, vol. 35, no. 9. pp. 1959–1970, 2011.

[7] K. Vytal and S. Hamann, "Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis.," *J. Cogn. Neurosci.*, vol. 22, no. 12, pp. 2864–2885, 2010.

[8] J. Panksepp, "At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain," *Brain and Cognition*, vol. 52, no. 1. pp. 4–14, 2003.

[9] R. J. Davidson, "Affective neuroscience and psychophysiology: Toward a synthesis," in *Psychophysiology*, 2003, vol. 40, no. 5, pp. 655–665.

[10] G. Gainotti, "Emotional behavior and hemispheric side of the lesion.," *Cortex.*, vol. 8, no. 1, pp. 41–55, 1972.

[11] R. G. Robinson, K. L. Kubos, L. B. Starr, K. Rao, and T. R. Price, "Mood disorders in stroke patients. Importance of location of lesion.," *Brain*, vol. 107 ( Pt 1), pp. 81–93, 1984.

[12] M. A. Safer and H. Leventhal, "Ear differences in evaluating emotional tones of voice and verbal content.," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 3, no. 1, pp. 75–82, 1977.

[13] P. Reuter-Lorenz and R. J. Davidson, "Differential contributions of the two cerebral hemispheres to the perception of happy and sad faces," *Neuropsychologia*, vol. 19, no. 4, pp. 609–613, 1981.

[14] E. Harmon-Jones, "Contributions from research on anger and cognitive dissonance to understanding the motivational functions of asymmetrical frontal brain activity," *Biological Psychology*, vol. 67, no. 1–2. pp. 51–76, 2004.

[15] Y. Meir-Hasson, S. Kinreich, I. Podlipsky, T. Hendler, and N. Intrator, "An EEG Finger-Print of fMRI deep regional activation," *Neuroimage*, 2013.

[16] R. J. Davidson, "What does the prefrontal cortex 'do' in affect: Perspectives on frontal EEG asymmetry research," *Biological Psychology*, vol. 67. pp. 219–233, 2004.

[17] E. Harmon-Jones, "Clarifying the emotive functions of asymmetrical frontal cortical activity," in *Psychophysiology*, 2003, vol. 40, no. 6, pp. 838–848.

[18] J. A. Coan and J. J. B. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion," *Biological Psychology*, vol. 67. pp. 7–49, 2004.

[19] J. A. Coan, J. J. Allen, and E. Harmon-Jones, "Voluntary facial expression and hemispheric asymmetry over the frontal cortex.," 2001.

[20] P. Ekman, R. J. Davidson, and W. V Friesen, "The Duchenne smile: emotional expression and brain physiology. II.," *J. Pers. Soc. Psychol.*, vol. 58, no. 2, pp. 342–353, 1990.

[21] R. J. Davidson and N. A. Fox, "Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants.," *Science*, vol. 218, no. 4578, pp. 1235–1237, 1982.

[22] G. Vecchiato, L. Astolfi, F. D. V. Fallani, F. Cincotti, D. Mattia, S. Salinari, R. Soranzo, and F. Babiloni, "Changes in brain activity during the observation of TV commercials by using EEG, GSR and HR measurements," *Brain Topogr.*, vol. 23, no. 2, pp. 165–179, 2010.

[23] G. Vecchiato, J. Toppi, L. Astolfi, F. De Vico Fallani, F. Cincotti, D. Mattia, F. Bez, and F. Babiloni, "Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements," *Med. Biol. Eng. Comput.*, vol. 49, no. 5, pp. 579–583, 2011.

[24] R. Jenke, A. Peer, and M. Buss, "Feature Extraction and Selection for Emotion Recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 327–339, Jul. 2014.

[25] R. Khosrowabadi, H. C. Quek, A. Wahab, and K. K. Ang, "EEG-based Emotion Recognition Using Self-Organizing Map for Boundary Detection," *2010 20th Int. Conf. Pattern Recognit.*, pp. 4242–4245, 2010.

[26] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis," *IEEE Trans. Affect. Comput.*, vol. 1, no. 2, pp. 81–97, 2010.

[27] S. Valenzi, T. Islam, P. Jurica, and A. Cichocki, "Individual Classification of Emotions Using EEG," *J. Biomed. Sci. …*, no. June, pp. 604–620, 2014.

[28] S. Koelstra, C. Mühl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, pp. 18–31, 2012.

[29] G. S. Taylor and C. Schmidt, "Empirical Evaluation of the Emotiv EPOC BCI Headset for the Detection of Mental Actions," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1. pp. 193–197, 2012.

[30] N. a Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur, "Validation of the Emotiv EPOC(®) EEG gaming system for measuring research quality auditory ERPs.," *PeerJ*, vol. 1, p. e38, 2013.

[31] R. N. Khushaba, L. Greenacre, S. Kodagoda, J. Louviere, S. Burke, and G. Dissanayake, "Choice modeling and the brain: A study on the Electroencephalogram (EEG) of preferences," *Expert Syst. Appl.*, vol. 39, no. 16, pp. 12378–12388, 2012.

[32] R. N. Khushaba, C. Wise, S. Kodagoda, J. Louviere, B. E. Kahn, and C. Townsend, "Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking," *Expert Syst. Appl.*, vol. 40, no. 9, pp. 3803–3812, 2013.

[33] Y. Liu, O. Sourina, and M. R. Hafiyyandi, "EEG-based emotion-adaptive advertising," in *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*, 2013, pp. 843–848.

[34] R. Ramirez and Z. Vamvakousis, "Detecting emotion from EEG signals using the Emotive Epoc device," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7670 LNAI, pp. 175–184.

[35] D. Ariely and G. S. Berns, "Neuromarketing: the hope and hype of neuroimaging in business.," *Nat. Rev. Neurosci.*, vol. 11, no. 4, pp. 284–292, 2010.

[36] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986.

[37] J. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," *Technical Report MSR-TR-98-14* . 1998.

[38] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Mach. Learn.*, vol. 59, no. 1–2, pp. 161–205, 2005.

[39] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, 2009.

[40] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis, and W. V Friesen, "Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology. I.," *J. Pers. Soc. Psychol.*, vol. 58, pp. 330–341, 1990.