Relevance of EEG Input Signals in the Augmented Human Reader

Inês Oliveira
CICANT, University Lusófona
Campo Grande, 376, 1749-024, Lisbon
PORTUGAL
ines.oliveira@ulusofona.pt

Ovidiu Grigore, Nuno Guimarães, Luís Duarte
LASIGE/FCUL
University of Lisbon, Campo Grande, 1749-016, Lisbon
PORTUGAL
{ogrigore | nmg}@di.fc.ul.pt

ABSTRACT
This paper studies the discrimination of electroencephalographic (EEG) signals based in their capacity to identify silent attentive visual *reading activities* versus *non reading* states.

The use of physiological signals is growing in the design of interactive systems due to their relevance in the improvement of the coupling between user states and application behavior.

Reading is pervasive in visual user interfaces. In previous work, we integrated EEG signals in prototypical applications, designed to analyze reading tasks. This work searches for signals that are most relevant for reading detection procedures. More specifically, this study determines which features, input signals, and frequency bands are more significant for discrimination between *reading* and *non-reading* classes. This optimization is critical for an efficient and real time implementation of EEG processing software components, a basic requirement for the future applications.

We use probabilistic similarity metrics, independent of the classification algorithm. All analyses are performed after determining the power spectrum density of delta, theta, alpha, beta and gamma rhythms. The results about the relevance of the input signals are validated with functional neurosciences knowledge.

The experiences have been performed in a conventional HCI lab, with non clinical EEG equipment and setup. This is an explicit and voluntary condition. We anticipate that future mobile and wireless EEG capture devices will allow this work to be generalized to common applications.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *user centered design, evaluation, interaction styles.*

General Terms
Design, Experimentation, Human Factors, Measurement

Keywords
Reading Detection, HCI, EEG Processing and Classification, Similarity Metrics, Feature Relevance Measurement.

AH'10, April 2-4, 2010, Megève, France.

1. INTRODUCTION
The understanding and use of human physical and physiological states in computational systems increases the coupling between the user and the application behavior. The integration of physiological signals in applications is relevant in the design of universally-accessible interactive systems and will become more relevant as new computing paradigms such as ubiquitous computing [7] and ambient intelligence [1],[14] develop.

The use of neurophysiological signals, and in particular electroencephalograms (EEG), has been widely reported in the context of an important example of coupled interaction systems: BCI's [4],[5],[16]. These interfaces exploring the information at its source, the brain. EEG signals are frequently chosen because of their small temporal resolution and non-invasiveness [9] and also due to its relative low cost capture device settings.

Visual user interfaces often require reading skills. The users' reading flow is highly influenced by their concentration and attention while interacting with applications. The application visual characteristics and users' cognitive state can decrease readability and degrade the interaction.

Augmented reading applications should adapt to the user's reading flow through the detection of *reading* and *non-reading* states. Reading flow analysis also improves the understanding of the users' cognitive state while interacting with the applications and improves the current empirical style of usability testing [9]. In previous work, we integrated EEG signals in two prototypical applications, designed to analyze and assist reading tasks. These applications are briefly described further down in this paper.

This paper focuses on the discrimination of EEG signals based in their relevance with respect to the identification of silent attentive *reading* versus *non reading* tasks, therefore finding the importance of each EEG signal for the reading detection procedure. The ultimate goal of this study is to allow a robust selection and weighting of input signals, which we deem critical for a feasible, efficient, and real time implementation of EEG processing software components, our augmentation approach.

EEG processing literature generally refers feature vectors of some extent. We have dealt with data dimensionality reduction in the processing pipeline by using Principal Component Analysis [9]. PCA does not consider the spatial distribution of the input signals nor the functional neurosciences knowledge. Neurosciences map cognitive processes into skull areas.. Quantifying the importance of each input signal in relation to reading detection will help verifying what electrodes and frequency bands are more involved in the reading cognitive process, and builds on the functional neurosciences knowledge.

The analysis of EEG signals relevance is performed after determining the power spectrum density (PSD) of delta, theta, alpha, beta, and gamma rhythms (the known EEG frequencies bands) in each of the captured EEG streams. We then apply probabilistic similarity measures [10], which are independent of the classification algorithm, in each of these streams to detect the main differences, and to discriminate between visual *reading and non reading* activities. All results obtained about the importance of the input signals are provided and crossed against functional neurosciences knowledge.

Our experiments were performed in a conventional HCI lab, with non clinical EEG capture equipment. This is not a limitation to overcome but rather a feature and an *apriori* requirement of our design. Even if the results can be further validated in clinical settings (*in vitro*), our goal is to address real life situations (*in vivo*) which have harsher stability, noise and artifact conditions.

We predict future mobile and wireless EEG capture devices will allow the generalization and extension of this work to common tools and applications. The broader goal of this work is to design and develop usable and robust software components for integration in interactive systems that reach higher adaptation levels through this augmentation approach.

2. EXPERIMENTAL SETTINGS

EEG signals were captured using MindSet-1000, a simple digital system for EEG mapping with 16 channels, connected to a PC using a SCSI-interface. These channels are connected through pure tin electrodes (sensors) to a cap made of elastic fabric, produced by Electro-Cap International.

Figure 1. MindSet-1000 and Electro-Cap Intl Cap.

Figure 2 shows the electrodes mapping that are used in our study. The EEG signals are amplified in differential manner relative to the ear electrodes and are sampled with 256Hz frequency. All requirements indicated by suppliers and technicians were fulfilled [9]. These included "grounding" the subjects and keeping the impedance in each electrode bellow 6000Ω, through the thorough application of conductive gel.

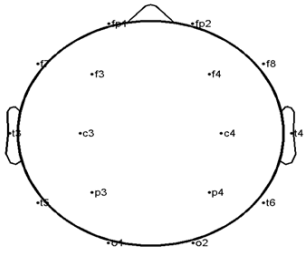


Figure 2. Mapping of used EEG electrodes (Int 10-20 method).

The first 5000ms and the last 3000ms of each trial are discarded for avoiding possible artifacts caused by start-end of the recording process. To assure the reliability of capture procedure the experiment was also tested using a professional medical capture device, in use in a Hospital, which setup was entirely prepared and tuned by expert technicians [9]. The results obtained with both capture devices were validated by an EEG specialist and a consistent set of sample results was produced.

2.1 Read and Not Read Experience

The capture experiments, object of the relevance analysis described in this paper, were based in the presentation of alternate blank and text screens containing about 40 lines of daily news text. The duration of such screens differed according with the ability to keep subjects concentrated in the task [9]. Text screens were presented in longer periods (30s) than black screens (20s). These types of periods were interlaced: one reading text sample, followed by 2 watch-only blank screens, and again back to read. All these periods were captured separately, allowing a small resting period, where the signal was not recorded.

Each capture trial included approximately 120s of both sample classes. All data was recorded without any previous or special training in a right handed female subject, mid thirties and without known vision disabilities (see discussion on this choice in the final section).

2.2 Assisted Reading Prototypes

In the context of these experiences, we designed simple prototype tools. *ReadingTester* tests in real time "reading event scripts", sequences of events with certain duration that are generated by the application. The subject is exposed to these events, and simultaneously the EEG is captured and analyzed. A detection performance report is built when the detection process stops.

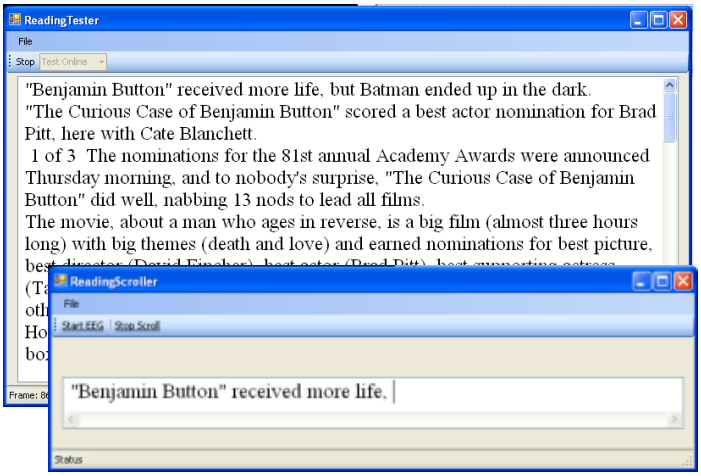


Figure 3. Assisted Reading Prototypes.

ReadingScroller aims at controlling the text scrolling through EEG signals: while the user is reading the scrolling should occur; if the user stops reading the scrolling should also stop. This is a trivial (from the functionality's viewpoint) Brain Computer Interface exploiting the reading detection capability, but with non trivial design challenges not reported here.

3. RELEVANCE ANALYSIS

Relevance analysis is performed after determining the PSD of delta (δ), theta (θ), alpha (α), beta1 (β1), and gamma (γ) rhythms in each of the 16 electrodes' input streams. This results in 16x5 PSD features streams, each with *reading* and *non-reading* samples. We then determined probabilistic dissimilarity measures separately in each of these streams, in order to quantify the dissimilarity between these two sample classes. The most relevant streams are those revealing larger significant differences between the *reading* and *non-reading* classes.

3.1 Probabilistic Dissimilarity Measures

Relative similarity is the relationship between two entities that share common characteristics with different degrees [10]. The larger it is, the greater the resemblance between the compared objects. *Relative dissimilarity*, on the other hand, focuses on the differences: the smaller it is, the greater the resemblance between the compared objects [10]. In our work, we compare the *dissimilarity* between *reading* and *non reading* samples sets. Both sets were approximated through Normal probability functions, since their samples result from discrete observations belonging to a large vector space.

Table 1 summarizes the probabilistic dissimilarity measures that were tested [2]. μ_i and Σ_i are respectively the mean vector and covariance matrix of the Normal distribution, noted as N_i, which approximates class i samples set. D_M is the squared Mahalanobis[1] distance between their means. In all the presented formulas, we assume $\Sigma_1 \neq \Sigma_2$.

For the sake of reproducibility of this work, the remainder of this section briefly describes each one of these measures.

3.1.1 Kullback-Leibler (KL) Divergence Based Measures

Kullback-Leibler Divergence is an asymmetric measure, a.k.a Relative Entropy or Information Gain. It quantifies, in bits, how close a distribution F_1 is from a (model) distribution F_2 [12], or, more precisely, the loss of information we incur in if we take F_1 instead of F_2 [8]. By definition, this measure between probability distributions $p_1(x)$ and $p_2(x)$ is determined by [10], [15],[8]:

$$KL(p_1, p_2) = \int p_1(x) \log \left(\frac{p_1(x)}{p_2(x)}\right)$$

For two Normal distributions $N_1(x)$ and $N_2(x)$ it becomes in the formula displayed in Table 1.

KL divergence cannot be considered a metric because it is asymmetric, that is: $KL(F_1, F_2) \neq KL(F_2, F_1)$ [12][8]. There are however measures such as J-Coefficient and Information Radius, which are symmetric versions of KL divergence.

J-Coefficient (JC) [2] is calculated by applying the KL formula symmetrically:

$$JC(F_1, F_2) = KL(F_1, F_2) + KL(F_2, F_1)$$

Information Radius (IR), also known as Jensen–Shannon divergence, is a smoothed symmetric version that is the average of the KL-distances to the average distribution [15]:

$$IR(F_1, F_2) = KL\left(F_1, \frac{F_1 + F_2}{2}\right) + KL\left(F_2, \frac{F_1 + F_2}{2}\right)$$

[1] $D_M^2(\mu_1, \mu_2; \Sigma) = (\mu_1 - \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2)$

Kullback-Leibler	$KL(N_1, N_2) = \frac{1}{2}\left(D_M^2(\mu_1, \mu_2; \Sigma_2) + tr(\Sigma_2^{-1}\Sigma_1 - I) + log\left(\frac{\Sigma_2}{\Sigma_1}\right)\right)$
J-Coefficient	$JC(N_1, N_2) = \frac{1}{2}\left(D_M^2(\mu_1, \mu_2; \Sigma_1) + D_M^2(\mu_1, \mu_2; \Sigma_2) + tr(\Sigma_1^{-1}\Sigma_2 - I) + tr(\Sigma_2^{-1}\Sigma_1 - I)\right)$
Information Radius	$IR(N_1, N_2) = \frac{1}{2}\log\left(\frac{\frac{1}{2}\det(\Sigma_1 + \Sigma_2) + \frac{1}{4}\lVert\mu_1 - \mu_2\rVert^2}{\det(\Sigma_2)^{\frac{1}{2}}\det(\Sigma_2)^{\frac{1}{2}}}\right)$
χ2 Divergence	$\chi^2(N_1, N_2) = \frac{\det(\Sigma_1\Sigma_2^{-1})}{2\,(\det(\Sigma_1\Sigma_2^{-1} - \mathrm{I}))^{\frac{1}{2}}} exp\left\{\frac{1}{2}\left(D_M^2(2\Sigma_2^{-1}\mu_2, \Sigma_1^{-1}\mu_1; 2\Sigma_2^{-1} - \Sigma_1^{-1}) + D_M^2(\mu_1, 0; \Sigma_1) - 2D_M^2(\mu_2, 0; \Sigma_2)\right)\right\} - 1$
Hellinger Coefficient	$HC^t(N_1, N_2) = exp\left\{-\frac{\mathrm{t(1-t)}}{2}\left(D_M^2(\mu_1, \mu_2; t\Sigma_2 + (1-t)\Sigma_1)\right)\right\} + \log\left(\frac{\det(t\Sigma_2 + (1-t)\Sigma_1)}{\det(\Sigma_2)^{\mathrm{t}}\det(\Sigma_1)^{1-\mathrm{t}}}\right)^{\frac{1}{2}}$
Chernoff Coefficient	$CC^t(N_1, N_2) = -\log\left(HC^t(N_1, N_2)\right)$
Bhatthacharyya Coefficient	$BC(N_1, N_2) = \frac{1}{8} D_M^2\left(\mu_1, \mu_2; \frac{1}{2}(\Sigma_1 + \Sigma_2)\right) + \frac{1}{2}\log\left(\frac{\frac{1}{2}\det(\Sigma_1 + \Sigma_2)}{(\det(\Sigma_1)\det(\Sigma_2))^{\frac{1}{2}}}\right)$
ℓ^2 Distance	$\ell^2(N_1, N_2) = \frac{1}{2\pi^{\frac{1}{2}}}\left(\frac{1}{\det(\Sigma_1)^{\frac{1}{2}}} + \frac{1}{\det(\Sigma_2)^{\frac{1}{2}}}\right) - \frac{2}{(2\pi)^{\frac{1}{2}}\det(\Sigma_1 + \Sigma_2)^{\frac{1}{2}}} \cdot \exp\left\{\frac{1}{2} D_M^2(\Sigma_2^{-1}\mu_2, \Sigma_1^{-1}\mu_1; \Sigma_2^{-1} + \Sigma_2^{-1}) - \sum_{i=1}^{2} D_M^2(\mu_i, 0; \Sigma_i)\right\}$

Table 1. Probabilistic dissimilarity measures.

3.1.2 χ^2 Divergence

χ^2 divergence is an asymmetric measure between probability distributions $p_1(x)$ and $p_2(x)$, *and is* determined by [8]:

$$\chi^2(p_1, p_2) = \int \frac{|p_1(x) - p_2(x)|^2}{p_1(x)}$$

The convergence in χ2 divergence implies convergence in KL-divergence, but the converse is not true [8]. This is because χ^2 divergence is strictly topologically stronger then KL-divergence, since $KL(P,Q) \leq \chi^2(P,Q)$.

3.1.3 Hellinger Coefficient (HC) Based Measures

Hellinger Coefficient (HC) of order *t* is a similarity measure between probability distributions $p_1(x)$ and $p_2(x)$, defined in [8]:

$$\mathrm{HC}^{\mathrm{t}}(p_1, p_2) = \int p_2(x)^t p_1(x)^{1-t}, 0 < t < 1$$

From this similarity-like measure, several dissimilarity coefficients have been derived. Chernoff coeficient (CC) of the order *t* is defined as [5]:

$$CC^t(F_1, F_2) = -\log\left(HC^t(F_1, F_2)\right)$$

This measure is related to KL divergence through its slope at *t*=0, it is smaller than KL divergence and it is less sensitive than the KL-divergence to outlier values [8].

There is also a special case symmetric metric for t=1/2, named Bhattacharyya Coefficient (BC), defined by [10]:

$$BC(F_1, F_2) = CC^{\frac{1}{2}}(F_1, F_2))$$

BC measures the amount of overlap between two probability distributions.

3.1.4 Minkowski's Based Measures

The Minkowski's L_p distance with p ={1,2,3, …} defined in [2][5]:

$$MD^p(p_1, p_2) = \int |p_1(x) - p_2(x)|^p$$

All Minkowski measures are symmetric and differ only in the way they amplify the effect of outlier values. Minkowski's distances of first and second order, L_1 and L_2 distances, are also known as Manhattan and Euclidean distance respectively.

L_2 measure is defined by [10]:

$$\ell^2(p_1, p_2) = MD^2(p_1, p_2)$$

It defines the distance between two points in a Euclidean n-space– a real coordinate space with n real coordinates, in this case our samples.

3.2 Relevance Measurement Method

We assume that relevance is directly proportional to the differences determined by dissimilarity measures. So our procedure is based in the ordering of the 16x5 feature streams accordingly with the calculated dissimilarities.

The first step is applying all the dissimilarity measures to the feature streams. This results in 16x5 (80) real values, one for each stream, corresponding to the measured difference. In order to compare all these values, all streams are normalized and turned into percentages by applying the following formulas:

$$metric'(f) = \frac{metric(f) - \min_s(metric(s))}{\max_s(metric(s)) - \min_s\left(metric(s)\right)}$$

$$metric''(f) = \frac{metric'(f)}{\sum_s metric'(s)}$$

The first equation normalizes the range of each difference to the interval [0,1]. The second weights it in relation to the overall results obtained with the measure.

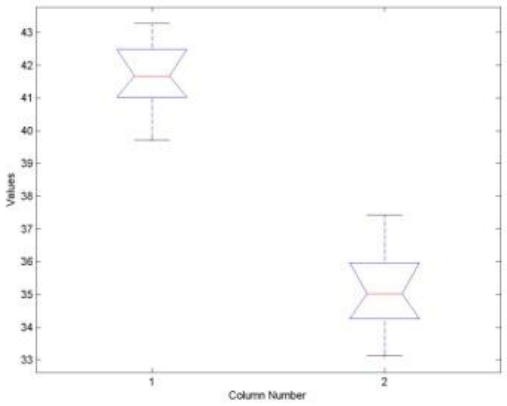
At this stage, after observing all the produced graphics, the Minkowski's Based Measures were discarded. These measures showed results that were too divergent from the ones provided by the rest of the metrics.

The final 16x5 weights, which we are using to quantify the importance of each 16X5 stream, are the average of all the measures. These weights are then ranked from the minimum (1) to maximum (80) importance and these are the results to analyze in order to determine signal relevance.

3.3 ANOVA Analysis

To statistically validate our conclusions we performed Variance Analysis, also known as ANOVA. It analyses the variation present in our experiments by statistically testing whether the statistical parameters of our groups of measures (bands, electrodes, etc.) are consistent, assuming that the sampled populations are normally distributed. If, for instance, this consistency holds for two electrodes or bands, then we can safely consider them correctly ranked.

ANOVA results are put into a graphic or table (Figure 4). The center line in the graphic represents the mean of each group, the above and below polygon lines, show it's the mean +/- variance values and the line segments delimit the confidence interval.



	SS	DF	MS	F	P	Crit.F
Between Groups	167,1	1	167,1	107,3	6,02E-08	4,6
Within Groups	21,8	14	1,6			
Total	**188,8**	**15**				

Figure 4. ANOVA for left (1) vs. right (2) Hemispheres.

The main ANOVA formula is given by:

$$F = \frac{\sigma_{BG}^2}{\sigma_{WG}^2}$$

where the numerator is the variance between groups, and the denominator is the variance within groups:

$$\sigma_{BG}^2 = \frac{\#Observations \sum_{i=1}^{\#Groups}(\mu_i - \mu)^2}{\#Groups - 1}$$

$$\sigma_{WG}^2 = \frac{\sum_{i=1}^{\#Groups}\sum_{j=1}^{\#Observations}(x_{ij} - \mu_i)^2}{\#Groups.\#Observations - 1}$$

The numerator of these formulas is represented in Squared Sum (SS) column (in the Table of the above Figure 4), while the Degrees of Freedom (DF) column contains the denominator. Total row is the sum of the columns. Mean Squares (MS) column is SS/DF. Critical F is got from the F distribution table and P is the probability of Critical F being as great as F. In the case of the above values, F is much greater that critical F and P is a significantly low value, so we can state that the statistical parameters of our groups of measures are consistent

4. PROCESSING AND ANALYSIS FRAMEWORK

All the processing functionalities are encapsulated in EEGLib framework, an object oriented toolkit implemented in C++ and MatLab [9],[3]. This framework provides tools for feature extraction and classification and also components for data modeling, such as EEG streams, frames, and iterators.

EEGLib includes several common EEG feature extraction procedures, including wavelets, power spectrum density (PSD), Event Related Synchronization (ERS) and other statistical measures. In the work described, we are using the mean PSD in Delta (δ) – 1 to 4 Hz, Theta (θ) – 4 to 8 Hz, Alpha (α) – 8 to 13Hz, Beta1 (β1) – 13 to 17Hz, and Gamma (γ) – 51 to 99 Hz – rhythms in all 16 electrodes. The analysis thus considers feature vectors composed by 16x5 real values. Mean PSD is determined in 1000ms frames with an overlapping of 500ms.

Our framework also has tools that support various standard learning methods, including neural networks, K-Nearest Neighbors (KNN), Ada Boost and Support Vector Machines. We have tested all of these tools, but for simplicity, current reading processing procedures are using the KNN provided in SPRTOOL MATLAB Toolbox.

5. RESULTS AND DISCUSSION

This section presents and discusses the results of the relevance ordering of input signals and bands.

5.1 EEG Signals Relevance Ordering

The relevance measurements ranks of all bands were averaged for each electrode. Figure 5 below presents the average values determined in all samples sets. The y-axis represents the importance rank average of all features of each electrode. For instance, the average of all features relative to O1 electrode has an average rank of 60 in 80.

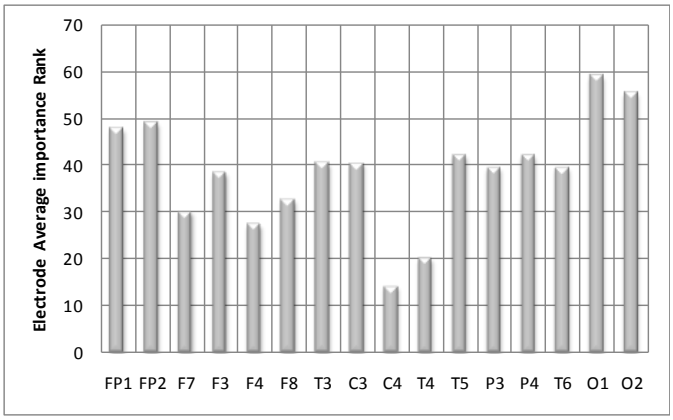


Figure 5. Average input signal relevance (ranks).

Figure 6 shows the locations of the most ranked electrodes. The electrodes that are not signaled have a rank inferior to 27.

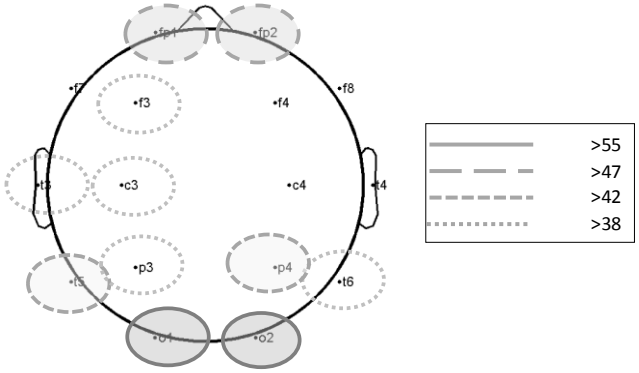


Figure 6. Average input signal relevance (locations).

It is clear that the main differences are dominant in the left hemisphere. This is in agreement with the study about reading tasks conducted by Bizas et al [2]. Their findings suggested that changes in PSD between reading tasks are restricted to left hemisphere. This hemisphere specialization is also confirmed by functional neurosciences experiments. It seems that about 90% of adult population has left-hemisphere dominance for language [13]. Broca's and Wernicke's regions, which are respectively responsible for speech and language understanding, are located in left hemisphere [13]. Wernicke's area influence is clear in our study: there is a visible importance elevation near electrodes T5, P3 and O1. We also expected a more visible influence of Broca's area in our results, near F7, C3 and T3, but this was inconclusive.

The highest ranked electrodes are in the frontal polar and occipital electrodes. Occipital lobe is where visual processing occurs [13], which supports our results regarding reading versus non-reading cognitive tasks. Frontal lobe is responsible for higher level processes, but we believe that it is more likely that the differences are due to eye artifacts.

5.2 Bands Relevance Ordering

The relevance measurement ranks of all electrodes were averaged for each band. Figure 5 shows the average results determined in all samples. The y-axis indicates the importance rank average of all features of each band. For instance, the average of all features relative to α band has an average rank of 50 in 80.

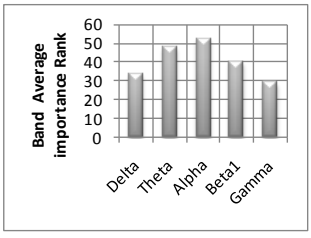


Figure 7. Average band relevance.

γ and δ bands are visibly less ranked that θ, α , and β1. We were expecting, due to previous related work, more relevant differences between these two groups of bands. γ rhythm is considered an important marker for attention [6]. It appears that visual presentation of attended words induces a γ rhythm in the major brain regions associated with reading and that this effect is

significantly attenuated for unattended words. A possible justification for poor γ band performance in our work can be that our experiments are focusing on reading instead of stressing attention. These results show the difference between visual reading and attention in the cortex activity.

Left hemisphere's δ, θ, and β1 rhythms were already used for differentiating reading tasks [13]. Significant differences in these rhythms between semantic tasks, the ultimate attentive reading activity, and visual, orthographic, and phonological tasks have been reported. The differentiation of θ rhythm was confirmed in our study, but this did not hold for δ and β1 bands. This cannot be due to the averaging effect, since these results are consistent with the non-averaged values (see next section).

The α rhythm, related with resting condition, demonstrated a good performance in our study, in spite of not being referenced in related work as the other bands are. Probably our non-reading task is behaving like a mental resting activity, when compared to reading task, thus causing the differentiation between the α bands of the sample sets of both classes.

5.3 Total Features Ordering

The importance measurement values were averaged for each of 16x5 features. Table 2 below displays the 10 highest average results, determined in all samples sets for each feature.

Rank	Average Relevance Rank	Electrode	Band
1	79,4	O1	Alpha
2	77,8	P3	Alpha
3	74,9	O1	Beta1
4	74,8	P3	Theta
5	73,9	O2	Alpha
6	73,5	O1	Theta
7	72,4	T5	Alpha
8	69,0	O2	Beta1
9	68,1	O2	Theta
10	66,3	T5	Theta

Table 2. 10 highest average feature relevance.

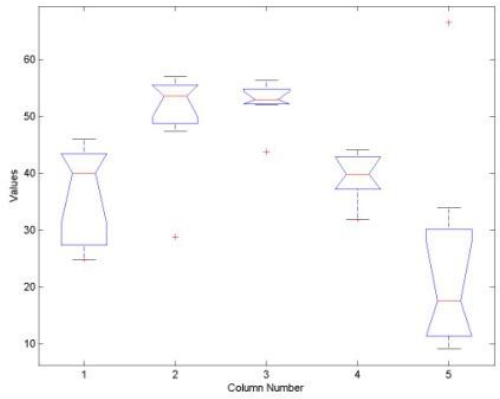
This ranking reinforces all the previous discussion, because all these values are located in the left hemisphere, and α and θ are the most frequent bands. It also shows that the averaging introduced in the previous analyses may minimize the importance of certain electrodes, namely P3 that appears twice in the top 10.

5.4 ANOVA Analysis Results

We performed several ANOVA test runs with different groups of measures, namely: left versus right hemisphere, skull areas, bands, electrodes, and features.

Erro! A origem da referência não foi encontrada. above (section 3) presented the ANOVA graphic and table for left and right hemispheres. These calculations were performed after averaging the ranks of all features related with each hemisphere. As we stated before, the results in the table indicate that the statistical parameters of the analyzed groups are consistent. This conclusion is reinforced by the graphic, which shows that the average ranks of both groups are statistically distinct with no possible overlap. We can also see that the left hemisphere importance is significantly higher than that of the right hemisphere.

The next figure shows the ANOVA result taking the the five tested bands as groups: δ, θ, α, β1, and γ (this order). These calculations were also performed after averaging the ranks in all features related with each band.



SV	SS	DF	MS	F	P	Crit.F
Between Groups	4158,5	4	1039,6	9,2	3,372E-05	2,6
Within Groups	3944,6	35	112,7			
Total	**8103,1**	**15**				

Figure 8. ANOVA for δ(1), θ(2), α(3), β1(4) and γ(5) bands.

The average ranks of θ and α were relatively higher and differentiated from the rest of the bands. Γ band performed poorly, showing the lowest rank and widest variation. According with the previous reasoning about ANOVA table results, we can also state that the statistical parameters of these groups are consistent, in spite of the F being close of its critical value.

To further detail this analysis we performed Multiple Comparisons: a technique that complements ANOVA and looks for specific significant differences between pairs of groups by checking the means among them. Figure 9 contains multiple comparison results for delta, theta, alpha, and beta1. Each line segment represents the comparison intervals of each group.

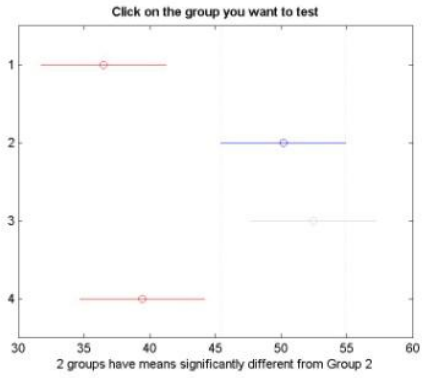
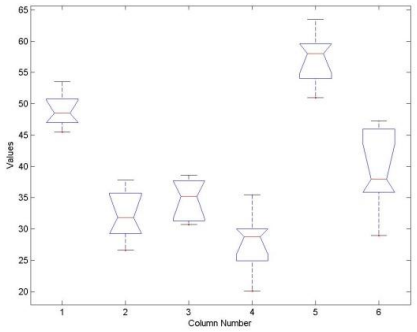


Figure 9. Multiple Comparison for δ(1), θ(2), α(3) and β1 (4).

δ and β1 bands comparison intervals were significantly different from the ones determined for θ and α rhythms. This also means that θ and α bands were significantly higher and distinct from the rest of the rhythms and, for this reason, they appear to be more relevant for classifying reading versus non reading tasks.

Figure 10 below displays the ANOVA result for specific skull areas: front polar, frontal, central, temporal, occipital, and parietal regions. These calculations were performed after averaging the ranks of all features related with each area.



SV	SS	DF	MS	F	P	Crit.F
Between Groups	4893,1	5	978,6	50,9	9,5E-17	2,4
Within Groups	807,4	42	19,2			
Total	**5700,5**	**47**				

Figure 10. ANOVA for Frontal Polar (1), frontal(2), central(3), temporal(4), occipital(5) and parietal(6) areas.

These groups' statistical parameters are also consistent, as the previous tables, since F is significantly higher than its critical value, and P is extremely small. Accordingly, with our previous results, we obtained average ranks relatively higher and distant from the remaining regions for the front polar and occipital areas.

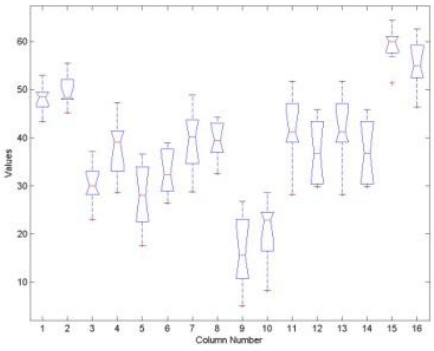
We then repeated the ANOVA process for all input signals using the average ranks in all features related with each electrode (see Figure 11). We did not discard any input signal at this stage in order to verify the averaging effect that we could get in the previous calculations. These results confirmed the previous discussion about areas. Front polar and occipital electrodes revealed higher ranks than the remaining electrodes, in spite of not being distant enough, especially front polar.

The values in the table also confirm these rankings as statistically consistent. F is once more greater that its critical value and P is very small. We then applied multiple comparisons to better analyze differences among electrodes (see Figure 12), and approximately got three groups, occipital, front polar and the remaining electrodes. Only for occipital electrodes, the comparison interval was significantly different from remaining electrodes group.

Finally, we applied ANOVA to individual features, but reducing its number to 16 by applying the previous conclusions (see Figure 13). Features were restricted to front-polar and occipital areas, and we also discarded the γ band.

The table supports that these rankings are statistically consistent, but we got here the lowest F value. However, F still is greater than its critical value and the probability of F being smaller than its critical value is very small (P).

δ band features from both occipital electrodes (9 and 13) worked poorly and showed a great variability. But the remaining features of these input signals were very concentrated and showed a relative distance regarding the rest of the groups. The variation of front polar related features (from 1 to 8) was more significant, especially for δ and β1 bands.



SV	SS	DF	MS	F	P	Crit.F
Between Groups	15849,8	15	1056,7	31,1	5,9E-33	1,8
Within Groups	3810,4	112	34,0			
Total	**19660,2**	**127**				

Figure 11. ANOVA for FP1(1), FP2(2), F7(3), F3(4), F4(5), F8(6), T3(7), C3(8), C4(9), T4(10), T5(11), P3(12), P4(13), T6(14), O1(15) and O2(16).

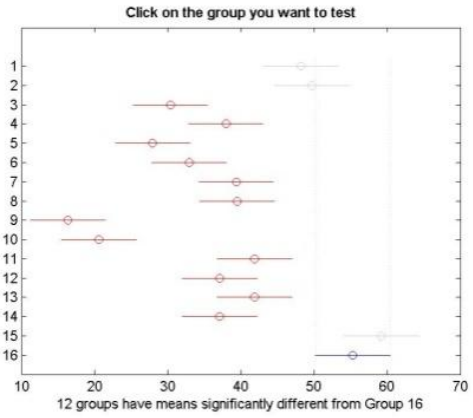
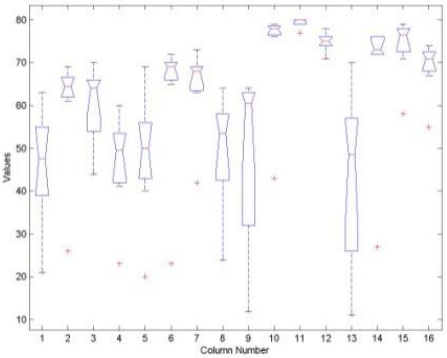


Figure 12. Multiple Comparison for FP1(1), FP2(2), F7(3), F3(4), F4(5), F8(6), T3(7), C3(8), C4(9), T4(10), T5(11), P3(12), P4(13), T6(14), O1(15) and O2(16).



SV	SS	DF	MS	F	P	Crit.F
Between Groups	17535,6	15	1169,0	6,7	4,59E-10	1,8
Within Groups	19584,3	112	174,9			
Total	37119,9	127				

Figure 13. ANOVA for FP1(1 to 4), FP2 (5 to 8), O1(9 to 12) and O2(13 to 16) with bands δ, θ, α and β1 respectively.

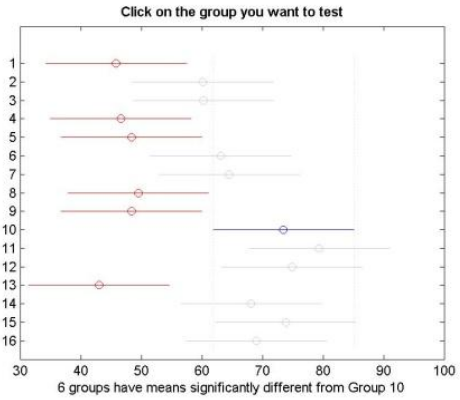


Figure 14. Multiple Comparison FP1(1 to 4), FP2 (5 to 8), O1(9 to 12) and O2(13 to 16) with bands δ, θ, α and β1 respectively.

Multiple comparisons (see Figure 14) revealed approximately three groups of comparison intervals: (I) 10 to 12 and 15, with higher ranks and significantly different from the next group; (II) 1, 4, 5,8,9 and 13, with lower ranks and significantly different from the previous group, and (III) Remaining features. Table 3 shows more detailed data about the first two groups.

Group	Average Relevance Rank	Electrode	Band
1	10	O1	Theta θ
	11	O1	Alpha α
	12	O1	Beta1 β1
	15	O2	Alpha α
2	1	FP1	Delta δ
	4	FP1	Beta1 β1
	5	FP2	Delta δ
	8	FP2	Beta1 β1
	9	O1	Delta δ
	13	O2	Delta δ

Table 3. Details of the two significantly different intervals.

Almost all **O1 electrode** comparison intervals were situated in the higher ranked group, revealing that this electrode appears to be consistently different between both reading and non-reading classes, since most of its bands were affected. δ band intervals were also consistently located in the lower ranked group, showing that it seems to be less relevant than the remaining rhythms.

6. CONCLUSIONS AND FUTURE WORK

This paper presented a study about the discrimination between the relevance of different types of EEG input signals with respect to their ability to identify silent attentive visual *reading* versus *non reading* cognitive tasks.

We have demonstrated that EEG input signals are not equally significant, and that we can quantify their contributions for the distinction between reading and non reading cognitive tasks. More than that, we outlined a systematic and quantitative method for relevance determination that can be applied to other cognitive tasks.

We presented results that reinforce that left hemisphere is dominant regarding reading tasks. We showed that its input signals consistently revealed higher dissimilarities between reading and non-reading samples than its homologues in right hemisphere. The results also indicated front polar and occipital areas, especially the latter, as also α and θ bands, related features as being more relevant that the remaining values. In opposition the some related work [12],[13], γ and δ bands results consistently performed poorly. In summary, we can state that:

For EEG-based silent reading detection, use mainly O1(θ,α,β1) and O2(α)

With this method, we can now proceed to the design of focused applications that exploit this significantly reduced set of human physiological features. The above specific conclusions are a first step towards the exploitation of this reduced set of signals in interactive applications targeted at assisted reading (such as *ReadingScroller*, briefly mentioned above). Having a reduced and optimized (for the cognitive task at hand) set of signals is a critical requirement for the optimization of the real time processing and for the use of the future light and portable EEG devices, where results are being reported that justify our expectations [11].

Our work elicits the following additional requirements and ideas that should be explored in sequence.

Calibration Procedures Design

Although our results were consistent with neurosciences knowledge and some of the existing related work, the presented analysis was performed with a single subject and a limited set of samples. This was a conscious choice in this stage to minimize the set of variables and tune the method. The repetition of the procedure with a larger number of subjects will now evaluate the degree of generalization of these results[2].

Our experience indicates that user differences will introduce some degree of diversity, such as skin conductance or hair type. In any case, differences are to be expected even when the subject is the same, due to biorhythmic cycle, sleepiness, or environmental conditions.

We aim to compensate these differences by designing adequate calibration procedures that adapt to the individual user profiles and conditions.

Dimensionality Reduction

As we said before, the ordering of EEG signals relevance, with respect to their ability to distinguish reading and non reading mental activities, is indispensable for the use of the future light and portable EEG devices. Signal ranking will allow the reduction of the number of sensors and turn the way users interact with augment reading applications more simple and natural.

In this context, we aim to include this knowledge in the current signal processing chain. A serious analysis about the impact of removing some of the less relevant features must be done. Reducing feature vector dimensionality will ultimately reduce processing time and allow the development of more effective real time applications.

[2] We referred above that around 90% of the population shows left hemisphere dominance for language [13].

Opportunities for Gamma Band Analysis

As we told before, γ rhythm is considered an important marker for attention [13]. However, it performed poorly in our study. Possible reasons for these results, in relation to the ones suggested in relative work, are the use of different type of features (PSD instead of wavelets) or distinct cognitive goals (reading versus non-reading instead of attentive versus non-attentive reading). A better understanding of this effect may be achieved through the use of wavelet coefficients for analyzing the γ band patterns in our experiments.

7. ACKNOWLEDGMENTS

This work was partially supported by Fundação para a Ciência e Tecnologia (FCT), Portugal, Grant SFRH/BD/30681/2006 and Ciência 2007 Program.

8. REFERENCES

[1] Aarts, E., Encarnação, J., *True Visions, The Emergence of Ambient Intelligence, Springer,* 2006.

[2] Bizas, E., Simos, G., Stam, C.J., Arvanitis, S., Terzakis, D., Micheloyannis, S. *EEG Correlates of Cerebral Engagement in Reading Tasks*, Brain Topography, Vol. 12, 1999.

[3] Oliveira, I., Lopes, R., Guimarães, N. M., Development of a Biosignals Framework for Usability Analysis (Short Paper), *ACM SAC´09 HCI Track*, 2009.

[4] Wolpaw, J. R. et al., "Brain–Computer Interface Technology: A Review of the First International Meeting", *IEEE Transactions on Rehabilitation Engineering*, Vol. 8, 2000.

[5] Millán, J.R., "Adaptative Brain Interfaces", *Communications of the ACM*, 2003.

[6] Jung, J., Mainy,N., Kahane,P., Minotti, L., Hoffmann, D., Bertrand, O., Lachaux, J., " The Neural Bases of Attentive Reading", *Human Brain Mapping*, Vol. 29, Issue 10, pp. 1193 – 1206, 2008.

[7] Krumm, J. (ed) (2010) *Ubiquitous Computing Fundamentals*, CRC Press, 2010

[8] Malerba, D. Esposito, F., Monopoli, M. , Comparing dissimilarity measures for probabilistic symbolic objects, *Data Mining III, Series Management Information Systems*, WIT Press, Vol. 6, pp. 31-40, 2002.

[9] Oliveira, I., Grigore, O. and Guimarães, N., Reading detection based on electroencephalogram processing, *Proceedings of the WSEAS 13th international conference on Computers*, Rhodes, Greece, 2009.

[10] Pekalska, E., Duin, R., "The Dissimilarity Representation for Pattern Recognition: Foundations And Applications", *Machine Perception and Artificial Intelligence*, World Scientific Publishing Company, Ch. 5, pp 215-254, 2005.

[11] Popescu, F. Siamac, F., Badower, Y., Blankertz, B., Mu ller, K., "Single Trial Classification of Motor Imagination Using 6 Dry EEG Electrodes". PLoS , ONE 2(7): e637, 2007.

[12] Shlens, J., "Notes on Kullback-Leibler Divergence and Likelihood Theory", *Systems Neurobiology Laboratory*, Salk Institute for Biological Studies, La Jolla, CA 92037, 2007.

[13] Steinberg, R. J., *Cognitive Psychology*, Thomson Wandsworth, 2003.

[14] Streitz, N., Kameas, A., Mavromatti, I., *The Disappearing Computer: Interaction Design, System Infrastructures and Applications for Smart Environments*, Springer, 2007

[15] Topsøe , F., *Jensen-Shannon Divergence and norm-based measures of Discrimination and Variation*, Technical report, Department of Mathematics, University of Copenhagen, 2003.

[16] Z.A. Keirn, J. I. Aunon, "A New Mode of Communication between Man and His Surroundings", *IEEE Transactions on Biomedical Engineering*, Vol. 37, 1990.