Evaluating Visualizations of Stereotactic EEG in Anatomical Context: an Experiment

Jean-Louis Dillenseger[1], Beatriz Sousa Santos[2,3], Pedro Teixeira[2], José Nunes[4], Carlos Ferreira[5,6]

1- Laboratoire de Traitement du Signal e de L'image, Univ. Rennes I, France
2- Dept. Electrónica e Telecomunicações, Univ. Aveiro, Portugal
3- Instituto de Engenharia Electrónica e Telemática de Aveiro
4- Dept. Comunicação e Arte, Univ. Aveiro, Portugal
5- Secção Autónoma de Engenharia e Gestão Industrial, Univ. Aveiro, Portugal
6- Centro de Investigação Operacional, Univ. de Lisboa, Portugal

***Resumo* – A Electroencefalografia Estereotáctica (Stereotactic Electroencephalogram SEEG) tem sido utilizada na definição de zonas epileptogénicas do cérebro. Uma solução possível para a análise destes sinais consiste na sua apresentação num contexto anatómico.**

Este artigo descreve uma experiência, realizada com observadores humanos, com o objectivo de averiguar qual de três esquemas de codificação da magnitude do sinal de SEEG veicula melhor este valor quando representado por esferas em contexto anatómico.

***Abstract* - Stereoeletroencephalography (SEEG) has been used to define and understand the organization of epileptogenic zones of the brain. The fusion of the SEEG signals and the anatomy on a common referential is a possible method for the analysis of these signals.**

This work describes an experiment conducted with human observers in order to evaluate three different coding schemes used to visualize the magnitude of SEEG signal in anatomical context.

I. INTRODUCTION

Epilepsy is the result of abnormal brain electric activities that mainly appear as synchronous (paroxystic) discharges within large populations of neurones belonging to brain structures implied during seizures. Investigation methods used in epileptology are aimed at defining and understanding the organization of the epileptogenic zone (from the areas originating the discharges to those secondarily affected by their propagation). Among these methods, Stereoeletroencephalography (SEEG) provides signals recorded with intra-cerebral electrodes. These signals bring major information on the dynamics of processes inside the brain structures. The visual analysis of SEEG signals is aimed at understanding the spatio-temporal dynamics of epileptic processes. More precisely, it tries to exhibit: 1) the existence of one or more areas generating independent activities, 2) propagation of paroxystic activities from one cerebral structure to another and 3) phenomena of synchronization or time delays between activities from subsets of structures.

Figure 1- SEEG in anatomical context.

One solution for a spatio-temporal analysis can be given by the fusion of the signals and the anatomy on a same referential [1]. The goal of the present work will be the evaluation of that technique in regard to the objectives of the medical task. In a first step we will restrict our study to the evaluation of the technique for the quantitative visualization of the signals. The magnitude of the signals is encoded by graphical glyphs that have a direct impact on the perception of the values [2,3]. Our evaluation will be devoted to the influence of the glyphs on the understanding and the analysis of the signals.

In the next sections we will introduce the objectives, describe an experiment performed using three different coding schemes for representing the magnitude of the signals, present the obtained results and draw some conclusions.

II. OBJECTIVES

The experiment presented in this work is just the first one from a set of experiments concerned with the evaluation of the influence of the glyphs on the understanding and analysis of the SEEG signals in anatomical context (as shown in fig. 1). Its objectives are related to the evaluation of the performance of human observers in extracting quantitative information from the visual representation of the signal coded through a glyph (a sphere in this case), without the presence of the scale.

In order to restrain the effect of external factors, our evaluation must be very strongly delimited in time, space, visual variables and amount of information [2,3]:

- Time - the time variation of the signals is of great importance for the analysis; however this variation would hold most of the observer's attention (the human brain is highly specialized in motion understanding). Moreover, the time variable in animations is not perceived as linear. To avoid these facts, our evaluation is performed on static frames.
- Space - SEEG signals are displayed on the 3D location of their measurement points; the external anatomy and the depiction of the depth electrodes give the anatomical location of these points. But, as shown in fig. 1, the orientation of electrodes can induce some perspective or superposition artifacts; thus on our evaluation we use parallel projection and a viewing direction perpendicular to the electrodes (fig. 3 and 4).
- Visual variables - the chosen glyph is a sphere with a size proportional to the value to encode. The other visual variables remain neutral and constant (color, texture, form, etc.) throughout the experiment.
- Amount of information - this factor interferes in two aspects, on the number of signals displayed simultaneously and on the quantification of these signals. To limit the first factor, only a few spheres (4 at the most) are displayed on each evaluation step. To limit the second, we quantify the scale of the values to integers in the interval [0,10]; the coding scheme of the glyphs is thus quantified to 11 spheres.

We have chosen three different coding schemes for the magnitude of the SEEG signal. The coding scheme of the glyphs must state the same organization level as the values. The 3D characteristics of the sphere allows three types of proportionality between the magnitude of the signal and the size of the sphere (fig. 2):

#1 - magnitude $\propto$ Radius of the sphere (R);
#2 - magnitude $\propto$ Projected surface of the sphere (πR^2);
#3 - magnitude $\propto$ Volume of the sphere ($(4/3)\pi R^3$);
where R is the radius of the sphere (as displayed in fig. 2).

The range of the coding scheme is kept similar for all the three coding schemes: from 0 to a maximum radius (encoding value 10).

Through this experiment we intent to study specifically if any of these coding schemes allows better performances for the human observers using them, in the referred conditions, i.e., if any of these scales can be considered "better" than the others. We have considered that a coding scheme can be compared to the others in two aspects: accuracy and speed. A coding scheme will be considered better than the others if it allows a more accurate and/or faster "measure" of the magnitude of the signal in a statistically significant way.

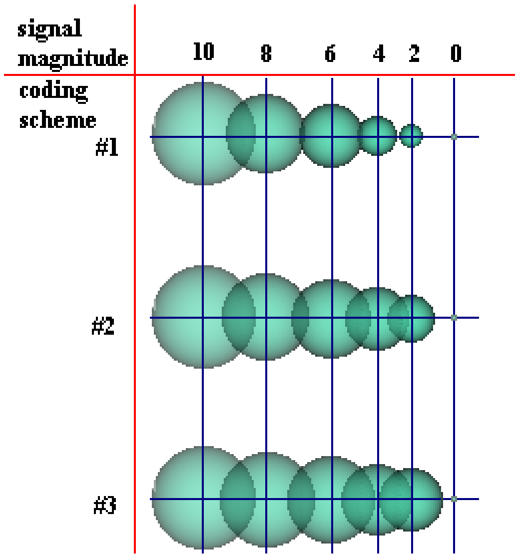

Figure 2- Three coding schemes used in the experiment.

III. THE EXPERIMENT

When we plan an experiment [4,5], it is necessary to think about the following aspects:

- The purpose of the experiment: what is being changed, what is being kept constant and what is being measured, i.e., the variables that should be used;
- A hypothesis, which needs to be stated in a way it can be statistically tested;
- What type and how many subjects should be used;
- What methodology should be used and how to compensate for possible biasing factors;
- What specific data should be collected and what statistical tests will be applied to the collected data.

Related to the first aspect, we had to choose the variables: the independent variables which the experimenter manipulates (i.e. the input), and the variables dependent upon them, the dependent variables (the output). The independent variables are expected to always remain uninfluenced by the dependent variables, but influence them. In our experiment only a few factors were manipulated so that the causal relationship between manipulations and observer performance was possible to establish.

Then, we had to formulate a clear hypothesis that predicted the expected performance effects associated with the experiment; this had to be stated in a way that could be tested statistically.

The choice of the subjects is related to the identified target population for which we want to draw some conclusions, they must be a representative sample of this population. The number of subjects is related to the statistical tests and the confidence degree we want to use.

The experimental methodology is also important; several types exist and its choice must be done bearing in mind the specific characteristics of the experiment and the advantages and disadvantages of each type.

In order to be able to obtain sound and well-founded results, all this process (and not only the choice of the statistical tests and collected data) was advised by a statistician. All the mentioned choices were also constrained by practical issues.

Finally, after defining and implementing the protocol and applying the experiment to the observers, we have performed a critical review of the experimental procedure, which allowed us to plan future experiments.

A. Variables and Hypothesis

In order to evaluate the accuracy and speed of the coding schemes we have chosen the following variables and hypothesis.

Independent variable:

coding scheme (#1, #2 and #3).

Dependent variables:

perceived magnitude error (real magnitude - perceived magnitude) and decision time (approximated by the reaction time).

Hypothesis:

As referred, we are interested in investigating if any of the three coding schemes is better than the others in two aspects (accuracy and speed). To reach a conclusion we have to state this in a way it can be tested statistically; this can be accomplished by first testing the equality of means and then, if it is rejected by testing which of the coding schemes is responsible for that rejection.

For the accuracy, we can say that, if the mean error in perceived magnitude is equal when using any scheme, then it is possible to consider that scheme #1 is as good as scheme #2 and as scheme #3 in accuracy. Statistically this can be stated as the following null hypothesis:

H_{0e}: $\mu_{e\#1} = \mu_{e\#2} = \mu_{e\#3}$ *versus* H_{1e}: $\mu_{e\#1} \neq \mu_{e\#2} \neq \mu_{e\#3}$

(where $\mu_{e\#n}$ is the mean error for coding scheme #n); the acceptation or rejection of this null hypothesis may be tested through a variety of statistical methods within the chosen confidence level.

Similarly for the speed, we can say that, if the mean error in perception time is equal when using any scheme, then it is possible to consider that scheme #1 is as good as scheme #2 and as scheme #3 in speed. This can also be stated as the following null hypothesis:

H_{0t}: $\mu_{t\#1} = \mu_{t\#2} = \mu_{t\#3}$ *versus* H_{1t}: $\mu_{t\#1} \neq \mu_{t\#2} \neq \mu_{t\#3}$

(where $\mu_{t\#n}$ is the mean perception time for coding scheme #n).

B. Subjects

One of the main goals of our experiment is to establish which coding scheme should be used to present a certain type of information (from SEEG signal) to medical doctors which need to used it. However, the experimentation presented here was mainly performed in order to establish a first basic experimental protocol and statistical methods. On the other hand, we intend also to study the more general issue of conveying quantitative information through this kind of visual coding. So, we have chosen subjects having a similar education and scientific levels. Thus, the 40 subjects that have participated in the experiment were Ph.Ds or pos-graduated students in engineering or sciences.

C. Experimental methodology

A repeated measures design was used [4]; in this type of design all the subjects appear in all the experimental conditions. There are no problems of subject allocation, however we must be careful with the order in which the subjects perform in these conditions.

In our case three experimental conditions exist, corresponding to the tasks performed for the three coding schemes. As experimental methodology, a within-group methodology was chosen [5,6]; the independent variable coding scheme was placed within-groups and all the observers performed the same task under each condition, i.e., observers are asked to evaluate the magnitude of spheres in several images and using all the coding schemes. In order to compensate for possible influence of certain side variables (as learning effects and interference due to the use of different schemes, nervous behavior in the first task or fatigue in the last task), the six possible sequences of tasks were used. This means that some observers first observed images using scheme #1, then #2 and finally #3; other observers have observed images using the sequence #2, #3, #1 and so on. The assignment of sequences to observers was performed randomly. For each coding scheme three different images were used (i.e. nine images were used for the complete experiment).

It would be desirable that all measured changes in the independent variables-output were due to an experimental effect (i.e., due to a change in the independent variable-input), however this usually is not the case; some changes in the output are brought about by other variables. The random assignment of sequences of tasks may decrease the effect of the previously referred side variables, however other variables can also have some influence (gender, age, profession, nationality, etc.). We were able to identify a few such variables (defining an observer profile) and have used them to characterize each subject.

In order to minimize the influence of some external conditions we chose to keep them from changing. Thus

- all the test images were generated in the same way, varying only the coding scheme;
- the point of view was the same for all the images;
- during the application of the experiment, the viewing conditions have been as similar as possible for all the observers (type of screen, viewing distance, ambient light, etc.).

D. Collected data

For each observer, the following variables were used to establish the observer profile:

Age (<25; [25,55]; >55), nationality, gender, familiarity with 3D video games or synthetic images; profession (medical doctor, or not); number of years of specialty and familiarity with 3D medical imaging (for doctors);

For each magnitude assessment, by an observer, the following variables were measured:

i) Main variables:

Error = real magnitude - perceived magnitude;

Perception time = time the observer takes to evaluate the magnitude (approximated by the reaction time);

ii) Variables concerning the images:

Coding scheme; sequence of coding schemes; number of spheres; spheres are occluding or not; spheres are contiguous or not; time from the onset of the experiment with this observer.

Variables concerning the spheres were collected in order to evaluate a possible influence of the number and type of spheres displayed on the same image on the perception of magnitude. Time from the onset of the experiment is expected to give some information on the effects of boredom or fatigue. Familiarity with 3D video games or synthetic images is collected in order to establish a possible non-formal training in 3D perception.

After performing the experiment, each observer was asked if he or she had any preferences about the coding schemes and why.

E. Protocol

A S/W package, for Windows platforms, was developed to allow an easy application of the protocol defined for the experiment. All observers received a simple explanation about the context and aims of the experiment and what was expected from them.

After they were fully informed and have agreed in proceeding, they were asked the information needed to define their profiles and then the protocol started.

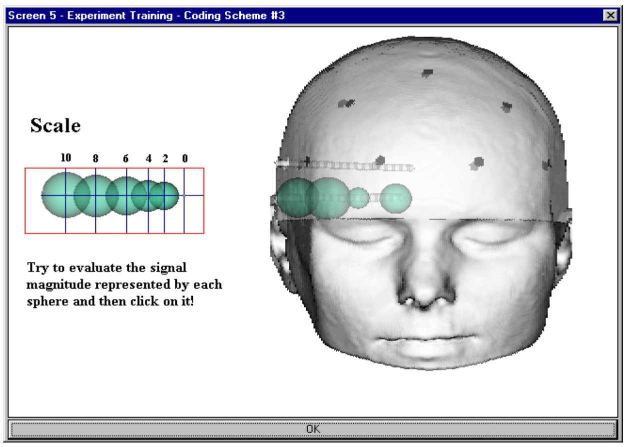

Figure 3- Interactive training for one of the coding schemes.

The protocol is divided in three similar parts (one for each coding scheme) where observers train themselves in using a certain coding scheme before observing all the images corresponding to same coding scheme. This training is interactive, consists of two screens as shown in fig. 3, and ends only when the observer decides he/she is ready to proceed.

After training for each coding scheme, the observers are shown three different images containing a certain number of spheres and are asked which is the magnitude of one of the spheres (as shown in fig. 4). The error in magnitude (real magnitude - perceived magnitude) and the reaction time are registered in a file (as well as most of the other variables referred in III.E).

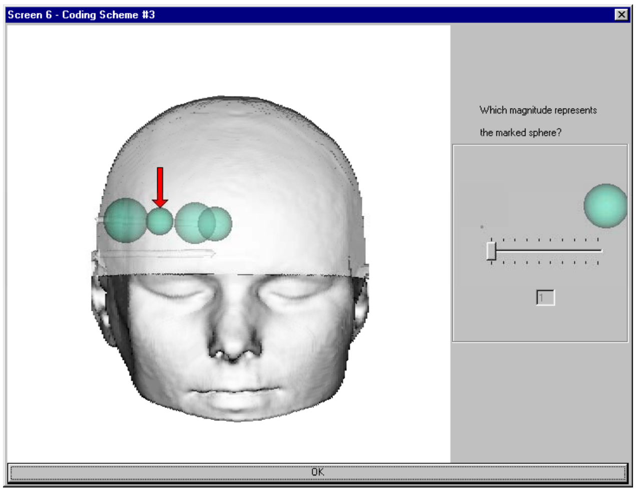

Figure 4- What is the magnitude represented by this sphere?

Before applying the experiment to the complete subject population, a pilot study was performed with 8 observers having a similar profile as the observer population. This study allowed testing the protocol and resulted in some minor adjustments.

IV. RESULTS

The results presented in this work were obtained applying a first experimental set to 40 human observers. These observers exhibited a relative homogeneous profile (Faculty staff and students; 25% of females). The aim of this first step was mainly to verify the protocol and the statistical analysis. Nevertheless, we also hoped that it would produce some results about the performances of observers using the three coding schemes, since the data obtained from such a number of observers can already be considered as statistically significant.

The application of the protocol to the 40 human observers yielded (along with the values of the variables characterizing each observer's profile) 120 error values (E_1, E_2, E_3) and 120 perception times (T_1, T_2, T_3) for each coding scheme:

E_n - (real magnitude – perceived magnitude) for coding scheme #n.

T_n - perception time for coding scheme #n;

Note that, $E_n \in Z$ and T_n is expressed in seconds.

The framework of the statistical analysis used on the obtained data was the following:

- The first analysis performed on these data was an Exploratory Data Analysis (EDA) [7], in order to get an overview of certain data characteristics (such as ranges, asymmetries, the existence of outliers). It allowed a preliminary comparison among the three coding schemes and helped on the choice of other statistical techniques to be used to further analyze the obtained data;
- Normality tests performed on the data sets (absolute error values and times) indicated that they could be characterized by their means and standard deviation;
- The hypothesis H_{0e} and H_{0t} (the equality among the three means corresponding to the absolute error values and times of all the coding schemes) was tested using a one way analysis of variance (ANOVA) [8]. This analysis is supposed to reveal if the differences between the data sets are due (or not) to the independent variables.
- Since the equality of the means hypothesis (both H_{0e} and H_{0t}) were rejected, a proportion (the Tuckey's pairwise comparisons method [8]) was used to establish which of those means were responsible for this rejection.
- Finally the observers preference to a specific coding scheme was taken in account.

All these statistical analyses were performed using the commercial s/w package MINITAB [9].

A. Overview of the Data

A well-known EDA technique, the box plot [7], was used. These plots display the maximum and minimum values as well as a central box indicating the location of the 50% central values. These are the values between the lower and upper quartile, in our case the 39th and the 90th elements from the sequence of 120 sample values in ascending order. Outliers are displayed as asterisks (*) and the median as a small square(▪) drawn on a line (fig. 5).

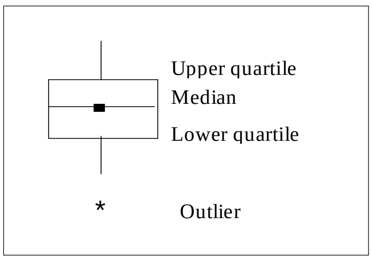

Figure 5- Box plot notice.

The box plots obtained for the variables E_1, E_2, E_3, T_1, T_2, T_3 are shown in fig. 6 and 7. Tables I and II show the median, minimum, maximum, quartile and range values for the same variables.

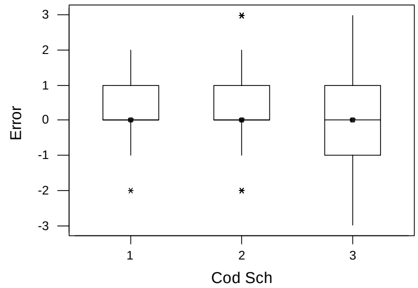

Figure 6- Box plot for the errors corresponding to coding schemes #1, #2 and #3 (E_1, E_2, E_3).

	Median	**Min**	**Max**	**Lower-Quart.**	**Upper-Quart.**	**Range**
E_1	0	-2	2	0	1	4
E_2	0	-2	3	0	1	5
E_3	0	-3	3	-1	1	6

Table I- Median, min., max., quartiles and range for the 120 error values (E_1, E_2, E_3) corresponding to coding schemes #1, #2 ,#3.

The errors corresponding to coding schemes #1 and #2 seem very similar. They have the same minimum, median and 50% of central values (between 0 and 1), the only difference being the existence of a maximum value (3) for the coding scheme #2 (which, however, is considered as an outlier). On the other hand, coding scheme #3, presents a larger range; it must be noted that the minimum values for coding schemes #1 and #2 are considered outliers (just one and two (-2) values, respectively) and this is not the case for coding scheme #3 (with seven (-2) values). Moreover, the 50% central values of coding scheme #3 are spread between –1 and 1.

Note that the median and the lower quartile values are coincident both for E_1 and E_2 (the 30th and the 60th values are zero), which is not the case for E_3. Analyzing the raw data, it can be verified that 66, 62 and 42 from 120 values for E_1, E_2 and E_3 respectively are equal to zero.

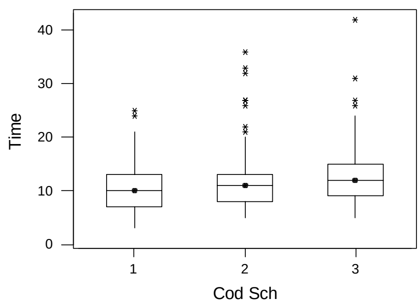

Figure 7- Box plot for the times corresponding to coding schemes #1, #2 and #3 (T_1, T_2, T_3).

	Median	Min	Max	Lower-Quart.	Upper-Quart.	Range
T_1	10	3	25	7	13	22
T_2	11	5	36	8	13	31
T_3	12	5	42	9	15	37

Table II- Median, min., max., quartiles and range for the 120 time values (T_1, T_2, T_3) corresponding to coding schemes #1, #2 and #3.

The minimum, maximum and median values of the perception times obtained for coding scheme #1 are smaller than for any of the other coding schemes; this is also the case for the 50% of central values. However the great difference observed in the ranges of the three coding schemes may be not so significant since the (worst) maximum values of coding schemes #2 and #3 are considered outliers.

B. Difference between schemes

After the EDA, which gave an overview of the data structure, the equality of means hypothesis, H_{0e} and H_{0t}, were tested using a one way ANALYSIS of variance, ANOVA[8]. This statistical method is similar to regression in that it is used to investigate and model the relationship between a response variable and one (or more) independent variable (mean error or time and coding scheme respectively, in our case). However ANOVA differs from regression in that no assumption is made about the nature of the relationship (i.e., the model does not include an explicit mathematical relation between variables). In effect analysis of variance extends the two sample *t*-test for testing the equality of two populations to a more general null hypothesis of comparing the equality of more than two means, versus them not all being equal. The *F* distribution is used and a unilateral test is performed. If the F_{obs} value (calculated from the data) is greater than the theoretical value $F(df_1,df_2)$ for a certain level of confidence, the null hypothesis is rejected.

In the case of the absolute values of the errors (as shown in the table of table III): $F_{obs} = 5,39 > F_{(2,117);0.95} = 3,09$; the null hypothesis ($H_{0e}$) was rejected for this confidence level (95%).

Source	df	SS	MS	F_{obs}
Cod Scheme	2	1,919	0,960	5,39
Error	117	20,812	0,178	
Total	119	22,731		

Table III- One way ANOVA to test the equality of means of the absolute errors corresponding to coding schemes #1, #2, #3 (E_1, E_2, E_3); with df: degrees of freedom; SS: sum of squares; MS: mean squares.

In order to establish which mean or means were responsible for this rejection, multiple comparisons of means was used (namely the Tuckey's pairwise comparisons method [8]) to conclude that means corresponding to coding schemes #1 and #3 were different.

Similarly, in the case of the perception times (as shown in the table IV): $F_{obs} = 3,43 > F_{(2,117);0.95} = 3,09$; the null hypothesis ($H_{0t}$) is also rejected for this confidence level (95%). The same method of multiple comparisons was used to conclude that means corresponding to coding schemes #1 and #3 are also different.

Source	df	SS	MS	F_{obs}
Cod Scheme	2	113,2	56,6	3,43
Error	117	1928,2	16,5	
Total	119	2041,4		

Table IV- One way ANOVA to test the equality of means of the times corresponding to coding schemes #1, #2 and #3 (T_1, T_2, T_3).; with df: degrees of freedom; SS: sum of squares; MS: mean squares.

C. Observers' preferences

Finally, the relation between the preferences expressed by the observers concerning coding scheme and their performances with the preferred coding scheme was investigated. Thirty-five from the 40 observers have expressed a preference by a certain coding scheme; from these, 20 have performed better when using the preferred coding scheme (the other 15 have not). The test of two proportions [8] was used to study if the proportion, p_1, of the observers who performed better using the preferred scheme was equal to the proportion, p_2, of the observers who did not. In other words, to investigate if there is an influence of the preference on the performance. The hypothesis were H_0: $p_1=p_2$ *versus* H_1:$p_1 \neq p_2$. Since Z_{obs} is 1,21 and $Z_{(0,975)}=1,96$, the null hypothesis H_0 is not rejected. This means that the proportion of best performances did not differ depending on the preference. However this result must be considered with caution, since due to the size of the sample, the normal approximation may be not very accurate.

V. DISCUSSION

After applying the experiment to the 40 observers, a critical review of the procedure and of the results were performed.

The review of the experimental procedure was based on the statistical results obtained and on the experience the experimenter herself had accumulated from observing the subjects performing the experiment and interviewing them after they have completed the tasks. This review considered the following issues:

- User preparation - the instructions given to the observers seemed adequate, as well as the amount of task practice they were allowed to obtain before performing those tasks;

- Structure of the tasks - some users have referred that the size and location of the spheres displayed on each image could influence the performance; this implies the need of a careful choice of these characteristics;
- Task duration - the duration of the tasks the observers were asked to perform seemed too small when compared to the preparation time, thus it seemed to be possible to increase significantly the length of the tasks (evaluation of more spheres), without the risk of producing fatigue or boredom in the observers. This will have the advantage of producing a larger data set that will allow to perform other statistical analysis, or the same with a greater confidence;
- Impact of independent variables - some of the users referred that the sequence in which the coding schemes were used could make a difference.

This review has been used to design another experiment that is currently being applied in order to clarify some issues that were left open by the experiment described in this work.

To establish exactly what has been found out by this experiment, the critical review of the results was performed and the following points were considered relevant:

- Size effect - the absolute size of the differences found in the dependent variables is important in assessing the results. From the point of the view of the errors (not taking into account the outliers identified by the statistical package used to produce the box plots), the difference between coding scheme #3 and the others seems rather significant also from a practical perspective. In figure 6 we can observe a total range of 6 for coding scheme #3, versus 3 for the others; we can also notice that 50% of the errors made with coding schemes #1 and #2 are 0 or 1, whereas for coding scheme #3 are –1, 0 or 1. This means that observers not only make more errors with coding scheme #3 but also can make much larger errors; this can be important. Regarding to perception times, the situation looks different, since in a real scenario, the fact that users take more time to decide (in the range of tens of seconds) possibly will not make much difference.
- Alternative interpretations - perhaps the result obtained indicating that one of the coding schemes is less accurate than the others could be influenced by several other variables. This will have to be clarified in further experiments.
- Consistency between dependent variables - the results seem consistent in that one of the coding schemes was found worst both on accuracy and speed. However some inconsistency may be related to the fact that many observers clearly preferred coding scheme #1, that was not found as significantly different from coding scheme #2. Perhaps this preference is due to the specific way coding schemes are displayed to the observers (as shown is figure 2), in fact the envelop of the spheres corresponding to the preferred coding scheme is a straight line and this was immediately recognized by many of the observers.

The fact that coding scheme #1 was not discriminated from coding scheme #2, through the previous statistical analysis, could also be explained by the following alternative interpretations:

- The small sample of observations - in this first test, we reduced the length of the experimentation to 3 images per coding schemes in order to avoid a saturation of the observers. After this first evaluation we decided to enhance our protocol to 10 test images per coding schemes. This extension is still well supported by the first observers of the new experimental set.
- The observers' profile - this first test was conducted on a relatively homogeneous population (University teachers and students). No physicians have been involved in this test. The new experimental set will include more different profile of population and in particular neurologist.
- The number of independent variables – as the 3D view used or the information contained in the observers' profile (gender, familiarity with 3D images and profession) that we did not use, for the moment. For instance, a first informal test showed that there are a difference in results between the male and female groups whatever the coding scheme (higher accuracy but slower reaction time for the female). This informal test suggests that other independent variables could influence the results and have to be taken into account for the new experimentation.

Informally, just on looking on the coding schemes, most of the observers declared to prefer coding scheme #1. On the contrary, Bertin [3] suggests that the observer is more sensible to surfaces for the extraction of quantitative data. We hope that a new experiment using basically the same protocol, but performed with a greater number of observers having different user profiles, will produce new answers to that question.

VI. Conclusions

The experiment described on this work was only the first one from a set of experiments concerned with the evaluation of the influence of glyphs on the understanding and analysis of SEEG signals visualized in anatomical context. The obtained results seem to indicate that a slight advantage of coding scheme #1 over coding scheme #3, on accuracy and speed, could exist. However, due to the small sample used and to the fact that the observers' profiles are similar (mostly engineering students and teachers), no definite conclusions can be drawn from this study. Nevertheless it allowed to establish a basic protocol and provided several important clues on how to perform further experiments and analyze the obtained results.

An experiment using basically the same protocol will be performed with a greater number of observers having different user profiles, including doctors and observers with formal 3D training. The results obtained from this experiment are expected to allow the confirmation (or not) of the results obtained until now and establish relationships among some of the characteristics of the observers profiles and the performances with some of the coding schemes.

REFERENCES

[1] Rocha C., Dillenseger, J. L., Coatrieux, J. L., "Multi-Array EEG Signals Mapped with 3Dimensional Images for Clinical Epilepsy Studies", *Visualization in Biomedical Computing*, Hamburg, Sept. 1996, pp.467-476

[2] Tufte, E., *Visual Display of Quantitative Information*, Graphics Press, 1987

[3] Bertin, J., *Sémiologie Graphic, les diagrammes, les resaux, les cartes,* 3me ed., Impression des Editions de l'Ecole des Hautes Études en Sciences Sociales, 1998

[4] Dix, A., J. Finley, G. Abowd, B. Russell, *Human Computer Interaction*, 2nd. Ed. Prentice Hall, 1998

[5] Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland, T. Carey, *Human Computer Interaction*, Addison Wesley, 1994

[6] Shneidermen, B., *Designing the User Interface, Strategies for Effective Human-Computer Interaction,* 3rd ed., Addison Wesley, 1998

[7] Hoaglin, D., F. Mosteller, J. Tukey, *Understanding Robust and Exploratory Data Analysis*, John Wiley and Sons, 1983

[8] Freund, R., W. Wilson, *Statistical Methods*, Academic Press, 1993

[9] *User's Guide 2: Data analysis and Quality Tools*, MINITAB Inc., Statistical Software, 1998