Gestalt and Computational Perceptual Approach

Brain responses tendencies given by visual and auditory basic stimuli

Bruno Giesteira, João Travassos, Diamantino Freitas
Electrotecnic and Computer Engineering Department
University of Porto
Porto, Portugal
bgiesteira@fba.up.pt, jpctravassos@gmail.com, dfreitas@fe.up.pt

Diana Tavares
Neurophysiology Course
IPP: ESTSP-CEMAH
Porto, Portugal
tavares.diana@gmail.com

GUI interfaces require considerable visual attention for their operation excluding the access to important information coded only in the layout. In an Era of mobile devices, we must enhance the auditory designs, to facilitate the interactive contents access to the blind, people with low vision, and/or in any use context. This essay is part of an experimental approach at the human perception based on the theories of form - Gestalt - and the Computational in order to process and implement the brain acquisition signal, obtaining relations between the visual and sound stimuli. We present a computational approach that underlay the electrical signal acquisition of the brain to stimuli response – "Event-Related Potentials" (P300) – based on a fundamental visual syntax that assumes the Gestalt phenomenology with new statistical interim results to the modeling multi-perceptive of information processing (visual and auditory), with the ultimate goal of framing a lexicon and/or basic patterns common that can be applied directly to a well-grounded development of GUI – "Graphic User Interfaces" and AUI – "Auditory User Interfaces".

Keywords - Perception; Event-Related Potentials; Gestalt; Computational Theory; GUI; AUI

I. INTRODUCTION

Since 2008 the Signals and Systems Laboratory, is leading a new approach in the perceptual field in order to recognize correlations or tendencies between brain responses elicited by two different stimuli modalities, namely visual and auditory [1] whose the main goal is to enhance the interaction multimodality in GUI – "Graphic User Interfaces" and AUI – "Auditory User Interfaces". Statistical data are presented in order to guide future development of auditory icons [2] and "hearcons" [3].

In the essay, first we present the "Research Fundamentals" where it explicit our main motivations and the state of the art regarding the perceptual theories and brain signal acquisition (Event-Related Potentials). Then we explained our laboratorial "Methodology" particularly regarding to the stimuli used and the brain acquisition signal, and then we organized the statistical interim results in the topic "Conclusions and Future Work" that correlate visual and auditory stimuli regarding the velocity of brain recognition (m/s) as well its energy/resources to process the task. (m/v).

II. RESEARCH FUNDAMENTALS

A. *Pleas and Motivations*

Beyond the neurophysiologic and computational approach we primarily faced a perceptual issue. In terms of perception, we excluded any narrow approach based on only one line/school dogma of theoretical thinking. However, we identify ourselves with the Gestalt phenomenology [4] and Marr's computational theory [5]. Completely different conceptions about visual perception but, in our view, do not render and even complement each other in a Top-Down perspective. Perhaps because the first rests on to descriptive generalizations that make sense and definitely contribute to the understanding and discussion sustained on the phenomenology of visual perception in the XX and XXI centuries [6], but are difficult to reproduce in scientific terms, and the second because it triggers for the first time procedures and scientific methodologies to explain and replicate the way the human mind processes visual stimuli, falling nevertheless in computational reductionism (possibly suitable to the area of Artificial Intelligence) that, putting aside the individuals' phenomenological consciousness negatively conditioned Marr's theory. Nevertheless, revolutionized the way we currently investigate the areas of perception and cognition. In an increasable operative and neurophysiologic perspective [7].

B. *Theory of Form – Gestalt*

Our research assumes, contrary to Marr's theory, the subjective nature of the stimuli by the direct influence of the individual conscience [4] [8] e.g., color or even dots, although isolated from a whole context, have subjective phenomenological dimensions inherent to the educational and cultural factors.

As well as the "Feature-Integration Theory of attention" it is assumed that the visual scene is initially encoded in a number of separable dimensions, such as color, orientation, spatial frequency, brightness and motion direction [9]. Any features presented in the same central of "fixation" of attention are combined to form a single object (Gestalt:

 ISBN: 978-1-61208-108-3

"Pragnanz"). This idea was inspired in part by Hubel and Wiesel studies [10] and others who provided evidence about "features" separated in the visual cortex in which each represents a different perceptual dimension such as color, orientation and movement. The organization of the elements is presumably carried out based on factors such as similarity, proximity, contiguity, direction and similarity [11]. These laws appear, however, to operate at a very early stage, presumably before the attention function and before the process responsible for the constancy of the properties of objects such as shape, size, brightness, and so on. It is believed that this is because the mechanisms of attention and constancy presuppose the prior existence of separate competitors' entities or objects about which they operate. The same authors [11] of the essay "Grouping based on phenomenal similarity of achromatic color" suggest that the organization is at an early stage, based on a new principle which they called uniform connectedness. Any features presented in the same central of "fixation" of attention are combined to form a single object (Gestalt: "Pragnanz"), suggesting that regions of uniform stimuli that are interrelated, such as dots, lines or large areas, are interpreted by the perceptual system as a unit. On the essay "Detection Signal Theory – STD" [12] Tanner and Swets, whose the main concern involved the measurement of the relations between quality and intensity of a physical stimulus (e.g., light intensity or frequency of a tone) and the perceptual experience caused by that stimulus, attested that perceptual experiences have a continuum of magnitudes that are produced either by noise or by events.

There is thus a remarkable set of neurophysiological evidence indicating that the grouping of objects/stimuli according to Gestalt exists at an early stage, as well as the recognition of certain fundamental characteristics as color, texture, movement, etc., which by their similarity or difference, are distinguished by a perceptual level (joining or separating into different perceptual organization) [11].

C. Computational Approach

As a neurobiologist and computer scientist, David Marr's works [5] led to a theoretical analysis of vision as a scientific problem proposing vision theories in several areas, including edge detection and perception of depth and shape. The distinction he made between algorithmic / representational / computational and achievement levels of analysis, guided the thinking of vision scientists since then. The levels of computational analysis relate to the objectives and purposes of the system under research. This analysis attempts to characterize, in the abstract, what the system is designed to do.

The algorithmic level is to specify an algorithm or procedure for carrying out the purpose specified by the computational level. Take vision as a computational problem has improved communication between disciplines such as psychophysics, neuroscience and computer science contributing to progress in these areas.

In our laboratory approach, for the first questions of Marr's computational model we found out about the importance of selecting, as visual stimuli to be tested, some of the key elements and basic concepts of visual communication [13] as dot, line, texture, color, scale, depth, movement, not only because the brain processes them differently (which still is the case, particularly for color, motion and depth) but because, in addition, also incorporate a basic visual syntax emphasizing precisely the Gestalt phenomenology that suggests the instinctive demand of the human being in perceive a whole with meaning – "Pragnanz" – in the most consistent, regular and simple way as possible, helping us to obtain correlations and/or trends between visual and sound stimuli conceptually similar and in the same context.

The fact that these are the minimum units perceived in any visual composition that by "Pragnanz" gives it a shape and Uno meaning is extremely relevant in the future correlation between visual and audio settings, constituting the basic units of visual communication capable of structuring more complex image and sound scenarios. Moreover, also being the first stage of the neurophysiologic journey, disparate and orthogonal in the activation of specialized cells of cerebral cortex (e.g., colors; silhouettes; movements; depth) responds to subsequent questions of Marr's computational theory. Namely:

a) A set of preliminary questions to be asked, e.g., Why is it important to perceive dots, lines and colors?; What is its importance to the individual and his relationship with the world?; Why should the system work to make these visual stimuli explicit?; How can these be represented symbolically in the brain?;

b) Developing an algorithm capable of structuring the phenomenon in a neurophysiologic way at the cognitive evoked potentials (P300) level;

c) Testing and implementing the efficiency and robustness of the algorithm developed and the understanding of the data acquired at the neurophysiologic and perceptual level.

D. ERP-Event-Related Potentials

The recording of ERP – Event-Related Potentials – is a non-invasive electrophysiological investigation method whose goal is to evaluate some of the high level characteristics of information processing in the central nervous system. Each psychological operation in turn involves a temporal activation/inhibition pattern of neurons in a certain brain area. The sum of synchronously generated and event-locked postsynaptic potentials is recorded at the scalp in a form of an ERP component – a potential deflection that is spatially localized and temporally confined [14].

The analysis of the ERPs has been reported as a significant contribute to the knowledge of neural processes that underlie highly specific skills in humans such as language processing, comprehension, visual analysis of faces, processing of emotional stimuli (affective processing) [15] [16], affective picture processing [17], attention, auditory discrimination [18], visual selective attention [19] and mere recognition of stimuli [20].

 ISBN: 978-1-61208-108-3

III. METHODOLOGY

A. Basic Stimuli

All visual stimuli used were created to be capable to translate objectively the fundamentals of the visual syntax: The three primary light colors - Red, Green, Blue - plus White, and the basic visual elements and concepts - Dot; Line; Texture; Depth/Dimension; Movement [13]. For the auditory stimuli we selected an audiological grammar used on clinical exams since that is scientifically accepted [21], translating to sound, as far as possible, the basic visual syntax with the sound parameters used to clinical purpose. For the translation of colors we used a fundamental note with 60dB in Si6 (B6) that has a frequency of 1.975 KHz. The Si6 was the note that stayed closer to the 2 KHz, a frequency usually used in clinical context [21]. The musical instrument used to provide the tones varied according to the color we wanted to translate and with the previous volunteers' correspondence. Therefore we choose the sounds of a classic guitar with nylon strings; a piano, a synthesizer; and a glockenspiel, all of them in Si6 (B6 - 1.975 KHz) with 60dB. To translate the visual concepts we used pure tones also with 60dB. A "beep" at 2KHz for the dot, a 2s sound at 2KHz for the line, a 2s tone burst for texture; a 2s sound of 2KHz of frequency varying in intensity progressively from 60dB to 34dB to translate depth/dimension; and a fundamental sound composed by an octave that initiate with the note Si6 of 2KHz to 20 KHz for the movement concept.

Please, consult the "reference" address to full data access, "in press" [22].

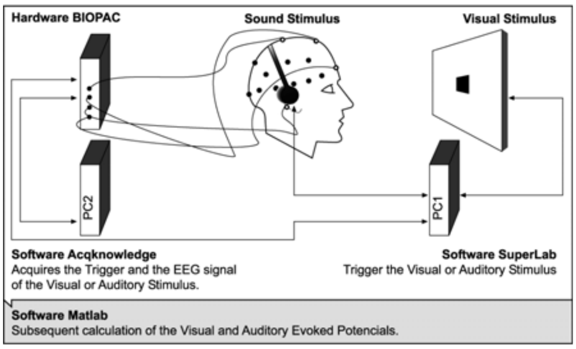

Figure 1. Acquisition Signal Scheme

B. Signal Acquisition

We divided the acquisition into three major parts and each part was preceded by a small interview. In the acquisition process the testing subject was instructed to discriminate a certain randomly appearing stimuli among other different randomly appearing stimuli by clicking on a button in his possession, while his brain activity was being recorded. The stimuli appeared one at a time, and we didn't cross stimuli modalities, i.e., we always separated visual stimuli from audiological stimuli (Figure 1). We always recorded a minimal of two times the brain activity for each stimulus the subject had to discriminate, to demonstrate reproducibility.

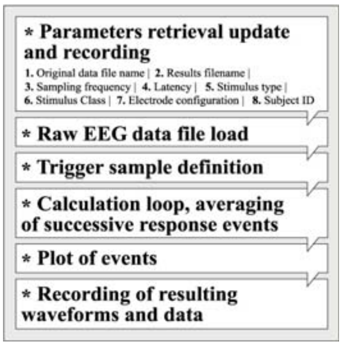

Figure 2. Algorithm Chart

All signals were processed and average techniques were applied to it, using "MatLab" software. We used "MatLab" software to develop the algorithm due to the fact that it is a more flexible tool and it allows us a deeper degree of analysis (Figure 2).

IV. CONCLUSIONS AND FUTURE WORK

Our laboratory study assumes the phenomenology of perception streamlined by Gestalt, founded upon a procedural computational methodology in which it developed an algorithm capable of calculating the Cognitive Evoked Potential (P300) through the acquisition of electroencephalographic signal of some audible and visual stimuli cited that with the same electrodes position, enables us to obtain the response time of the brain against the recognition of a specific stimulus - latency - as well as brain energy resources necessary for this purpose - amplitude. This is only possible because the sensory information from different modalities converge (between 200 and 300 ms after the stimulus) to areas of the cerebral cortex that integrate all information on poly-sensory events, i.e., all the visual, auditory, somatosensitive and olfactory information converge in associative multimodal areas located in the prefrontal, parietotemporal and limbic cortex.

Please, consult the "reference" "in press" [23] address to full data access. The images presented on the previous address represent the average (full line), and the standard deviation (dashed lines) of all acquisitions for each stimulus used. For now, with this approach we can sustain some preliminary results in a Statistical Report.

A. Statistical Report

The sample that here we presented consisted of 36 individuals, 10 (27.8%) of whom were male and the remaining 26 (72.2%) females. The average age of the sample was 22.75 years with a standard deviation of 5.699. The youngest person was 18 years old and the oldest 36 years. Statistical methods used:

a) Arithmetic Mean for the associations made by the volunteers as well as for ratings of the stimuli;

b) Confidence intervals to "catalog" each stimulus in terms of latency and amplitude;

c) Correlations to see which of the pairs have stronger correlations.

 ISBN: 978-1-61208-108-3

The complete statistical report could be consult in the following “reference” address [24].

B. Implications to GUI and AUI

On the bottom of this essay we present one example of correlation between one pair of stimuli which, like all the report data, will guide our team in future work to develop more sustainable and efficient auditory icons [2] and "hearcons" [3] thereby improving the speed of recognition - latency (m/s) - and the brain resources/energy (m/v) required to interact with a system in a multimodal way.

GUI interfaces require considerable visual attention for their operation. Providing to blind users only the textual contents of the web pages, excluding the access to important information coded in the layout of web pages, the same happened on mobile devices. If interfaces move also to the realm of auditory designs – AUI – these problems are mitigated.

REFERENCES

[1] S. Handel, “Perceptual Coherence: hearing and seeing”, 2006, 1st., Oxford University Press

[2] W. Gaver, “The SonicFinder: An interface that uses auditory icons”, Human-Computer Iinteraction, 1989, vol. 4, pp. 67-94, Lawrence Erlbaurn Associates, Inc.

[3] S. Brewster, P. Wright, and A. Edwards, “An evaluation of earcons for use in auditory human-computer interfaces”, Interchi'93, 24-29 April 1993, pp. 222-227

[4] W. Kohler, “Gestalt Psychology: The definitive statement of the gestalt theory”, 1992, Liveright

[5] D. Marr, “The Philosophy and the Approach”, in Steven Yantis - Visual Perception, 2001, pp. 104-123, New York: Psychology Press

[6] W. Metzger, “Laws of Seeing”, 2006, MIT Press

[7] S. Zeki, G. Watson, J. Lueck, J. Friston, C. Kennard, and J. Frackowiak, “A Direct Demonstration of Functional Specialization in Human Visual Cortex”, in Steven Yantis - Visual Perception, 2001, pp. 193-202, New York: Psychology Press

[8] L. Kaufman and I. Rock, “The Moon Illusion”, in Steven Yantis - Visual Perception, 2001, pp. 233-242, New York: Psychology Press

[9] A. Treisman and G. Gelade, "A Feature-Integration Theory of Attention", in Steven Yantis - Visual Perception, 2001, pp. 343-247, New York: Psychology Press

[10] D. Hubel and T. Wiesel, “Receptive Fields and Functional Architecture of Monkey Striate Cortex”, in Steven Yantis - Visual Perception, 2001, pp. 147-167, New York: Psychology Press

[11] I. Rock, R. Nijhawan, S. Palmer, and L. Tudor “Grouping Based On Phenomenal Similarity of Achromatic Color”, in Steven Yantis - Visual Perception, 2001, pp. 256-265, New York: Psychology Press

[12] P. Tanner and A. Swets, “A Decision-Making Theory of Visual Detection”, in Steven Yantis - Visual Perception, 2001, pp. 48-55, New York: Psychology Press

[13] D. Dondis, “La Sintaxis de la Imagen”, 1985, Gustavo Gili

[14] J. Kropotov, “Quantitative EEG, Event-Related Potentials and Neurotherapy”, 2008, pp. 253-291, Academic Press

[15] M. Eimer, “Event-related brain potential correlates of emotional face processing”, 2007, pp. 15-31, Neuropsychologia

[16] J. Petrek, “Pictorial cognitive task resolution and dynamics of event-related potentials”, 2008, pp. 223-230, Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub

[17] J. Olofsson, “Affective picture processing: An integrative review of ERP findings”, 2008, pp. 247-265, Biol Psychol

[18] J. Duarte, “P300- long-latency auditory evoked potential in normal hearing subjects: simultaneous recording value in Fz and Cz”, 2009, pp. 231-236, Braz J Otorhinolaryngol

[19] S. Hillard, “Event-related brain potentials in the study of visual selective attention”, 1998, pp. 781-787, Proc. Natl. Acad. Sci.

[20] E. Meijer, “The Contribution of Mere Recognition to the P300 Effect in a Concealed Information Test”, 2009, pp. 221-226, Appl Psychophysiol Biofeedback

[21] ASHA, “American Speech-Language-Hearing Association”, http://www.asha.org/default.htm (15.08.2010)

[22] www.giesteira.net/Stimuli, B. Giesteira, J. Travassos, D. Tavares and D. Freitas, “Brain’s electrical response to visual and auditory stimuli: Relations between the two stimuli modalities”, 2010, BMEI 2010, in press (15.08.2010)

[23] http://www.giesteira.net/ERP_Charts/, B. Giesteira, J. Travassos, D. Tavares and D. Freitas, “Brain’s electrical response to visual and auditory stimuli: Relations between the two stimuli modalities”, 2010, BMEI 2010, in press (15.08.2010)

[24] http://www.giesteira.net/Statistical_Report.pdf (15.08.2010)

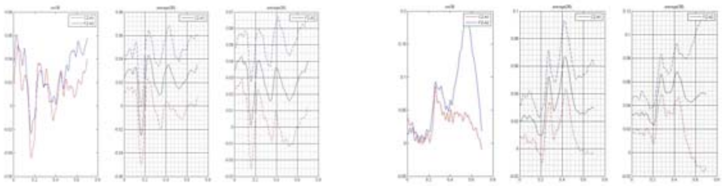

Figure 3. Average (full line), and the standard deviation of green color (left) and guitar (right) acquisition: http://www.giesteira.net/ERP_Charts/

TABLE I. P300 CORRELATION VISUAL AND AUDITORY STIMULI. SIGNIFICANCE LEVEL 0,05. PLEASE, CONSULT THE FOLLOWING ADDRESS TO FULL DATA ACCESS (STATISTICAL REPORT): HTTP://WWW.GIESTEIRA.NET/STATISTICAL_REPORT.PDF

	FZ				CZ			
Pairs	Latency (r)	Rate	Amplitude (r)	Rate	Latency (r)	Rate	Amplitude (r)	Rate
Green / Guitar	0,445	Moderate	0,116	Weak	0,302	Moderate	-0,002	Weak

TABLE II. “E.G.,” AVERAGE & STANDARD DEVIATION OF THE VISUAL AND AUDITORY STIMULUS

		FZ		CZ	
	N	***Latency***	***Amplitude***	***Latency***	***Amplitude***
Green	34	,3348971 (,02670193)	-,0104118 (,01911638)	,3404926 (,04178523)	-,0147663 (,01957021)
Guitar	34	,3235221 (,03180304)	-,0276206 (,01850645)	,3249632 (,03138339)	-,0280335 (,02009297)

ISBN: 978-1-61208-108-3