



The Visual Expression Process: bridging Vision and Data Visualization

José Fernando Rodrigues Jr., André G. R. Balan, Agma J. M. Traina, and Caetano Traina Jr. *

Universidade de São Paulo at São Carlos, Computer Science Department, Brazil

Abstract. Visual data analysis follows a sequence of steps derived from perceptual faculties that emanate from the human vision system. Firstly, pre-attentive phenomena determine a map of potential interesting objectives. Then, attentive selection concentrates on one element of a vocabulary of visual perceptions. Lastly, perceptions in working memory combine to long-term domain knowledge to support cognition. Following this process, we present a model that joins vision theory and visual data analysis aiming at settling a comprehension of why graphical presentations expand the human intellect, making us smarter.

Copyright Springer - The original publication is available at <http://www.springerlink.com/content/dq5q0703t3j52120/>

1 Introduction

Visual data analysis refers to the process of using visually mediated cognition in order to benefit from data; such cognition shall be driven intuitively and in elucidative manner. But how can vision mediate cognition? Why data visualization is potentially intuitive? In the next sections we put together a chain of concepts in order to delineate a coherent model that answers these questions.

The goal is to settle a perspective of visual data representations – as defined in Visual Analytics [15] – so that analytical insights can be achieved. Analytical assets are taken for granted as a necessary step in order to advance the mastery of any field of expertise. Abstract analysis allows for problems to be treated through smaller concepts leading to broader solutions and new hypotheses generation. In this line, here we address the needs pointed by Johnson *et al.* [6], who recommend the characterization of how and why visualizations work. Our broader goal is to develop principles for visual representations and, according to Card *et al* [3], such characterization is an initial step towards this goal. That is, first we must understand how visualizations enable cognition.

In Section 2 we build up a conceptual structure based on vision mechanisms. Following, in section 3, we trace the relation between vision and visualization. Section 4 carries a brief discussion. Section 5 concludes the paper.

2 Vision Mechanisms for Data Analysis

Neuroscience researchers state that, although the brain can perceive multiple visual targets simultaneously, it cannot process them in parallel. The solution is to restrict the consideration of the objects presented to the eyes. To do so, the mechanism of vision concentrates on small regions of the visual field, regarding single objects one after the other, a sequential process ruled by what is called *attention*.

* [junio,agrbalan,agma,caetano]@icmc.usp.br

2.1 Maps of Saliences

Attention solves the problem of overloading the brain, however attention casting is not so simple. Which regions/objects of the visual field should be considered as candidates for attention? In the early stages of vision processing, a set of characteristics found in a scene can potentially pop up to the eyes defining *maps of saliences*.

The idea of salience is to reinforce the neurons-mediated perception of the areas in the scene whose statistics of visual properties contrast with those of its surroundings. In Figure 1, we visualize the building of a map of saliences through an example; the figure demonstrates the framework proposed by Itti *et al.* [4].

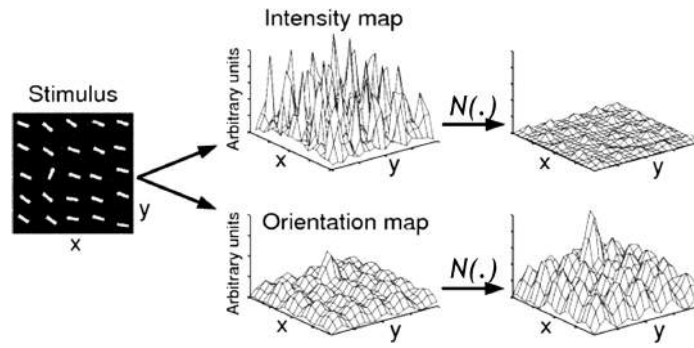


Fig. 1: The salience framework applied for color intensity and shape orientation. Illustration reproduced from Itti *et al.* [4].

After light stimuli reach the iconic memory (the primary short-term visual memory), they are processed in an early stage of the vision plexus. In parallel, this stage considers the whole visual field at multiple-scales of space and time. This first step produces a set of feature maps in the cortex area of the brain, each one summarizing how a particular visual feature is observed. For each map, a normalization operator $N(.)$ determines that several responses of similar amplitude generate maps without outstanding spots. At the same time, the normalization evidences the visual entities that contrast with their counterparts. The salience framework ends as it intersects all the generated maps onto a single panorama.

2.2 Attentive Selection

Once a map of candidates for attention is ready, it is necessary to choose one of them for processing. One at a time, the targets are considered according to the task to be performed (e.g., track all the red dots) or according to the data set (e.g., the blue dots pop up, what do they mean?). In the literature, the visual selection system has been widely assumed to have a pyramidal neuron structure. This model predicts a broad layer of neurons in its first level, narrowing down as it advances to upper layers, see Figure 2. The layers intercommunicate through feed-forward and feedback connections. According to Tsotsos [16], the pyramid of neurons performs three stages of processing:

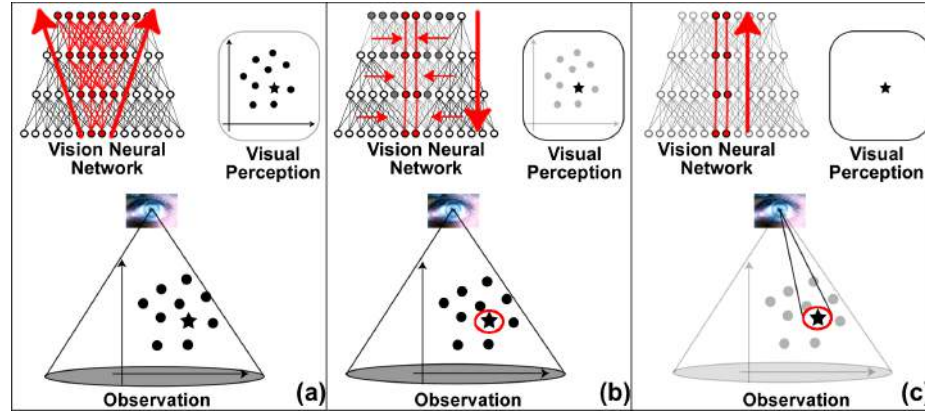


Fig. 2: Three-stage pyramidal visual selection.

- ◇ *bottom-up feed-forward*: the lowest layer receives stimuli from eye nerves, propagating the stimuli upward. The propagation generates a feed-forward inverted pyramid, see Figure 2(a). In this sub-pyramid, the activation of the connections and of the neurons depends on how the neurons are selective and biased. Biasing comes from upper layers according to the nature of the selection, task-oriented or data-oriented;
- ◇ *top-down feedback with WTA*: the upper layers return the initial stimuli through feed-back connections downward the pyramid, Figure 2(b). At each layer of this reflexive propagation, a WTA (winner-take-all) process takes place fine tuning the network operation. Initially, at the top layer, WTA considers the entire visual field to compute the neurons with largest response, the winners at the first layer. The winners, then, promote a WTA among their input neurons at the layer immediately below, identifying the winners at this lower layer. At the same time, the connections that do not contribute to the winners are pruned (attenuated). This three-way process – WTA, descend and prune – layer by layer, repeats down to the lowest (input) layer. At the end, the winner at the output layer is traced back to its perceptual origin at the input layer;
- ◇ *bottom-up straight path*: the stimulus originating at the target object propagates through the path of winner neurons and of active connections, Figure 2(c). Now, the propagation proceeds without distracting stimuli, on each receptive field, as if the target was projected on a blank background.

2.3 Cognition, Memory and Vision

After a target is focused, it is potentially useful for cognition. Cognition refers to the acquisition or use of knowledge [3], a process assisted by the memory system. The relation between memory and cognition is studied in works on *Cognitive Architectures*, such as ACT-R and Soar. A widely accepted cognition-memory relationship is enunciated by the *Soar* architecture: memory refers to mechanisms that maintain and provide access to information created or retrieved during the performance of a task. Any computationally complete system must support such functionality, because computation is

inherently a process that requires the temporary storage and manipulation of partial and intermediate products [19].

The *structural configuration* of memory, in Soar and other theories, more or less reflects the model described by Baddeley and Hitch [1]; working memory is comprised of three components: the central executive module, the phonological loop and the sketchpad. The central executive module determines the attention focus, guiding the visual perceptual system as, for example, by biasing the pyramidal selection mechanism. The phonological loop stores information related to sound, for example, a phone number that must be mentally repeated to prevent it from being forgotten. The sketchpad (also known as Visual Short-Term Memory – VSTM) is associated to the maps of saliences presented in section 2.1, storing information related to space and to visual features.

In the theory of working memory, VSTM is the main component in supporting cognition. Luck and Vogel affirm that people can retain only four to five single features in VSTM [10]. This is a problem because the greater the capacity of an individual's memory, the more information she/he has available for solving problems [7]. Hence, VSTM would limit cognition. However, according to Jiang *et al.* [5], the sense of vision is used as an extension of the working memory through VSTM. Even more, Jiang *et al.* argue that visual representations need not to carry details simply because one can always rely on the outside world as her/his visual memory. This fact is crucial not only for the interaction of human beings with the physical world, it also explains the importance of computer graphics for data analysis and reasoning. That is, knowing that memory, although limited, is the factor that supports cognition, then, graphical externalizations appear as a natural ally for cognitive abilities; this is because graphical externalizations expand memory capacity.

3 The Visual Expression Process

We have reviewed the mechanisms of maps of saliences, attentive selection and vision-supported cognition. Now, we use this knowledge to analyze how data visual representations work. We introduce a model that links vision and visual data analysis, a theme overlooked in the visualization literature. Our framework is named Visual Expression Process, it encompasses three components: pre-attentive stimuli, visual perceptions and interpretations.

3.1 Pre-attentive Stimuli

Pre-attentive stimuli are the constitutional features that impel maps of saliences. A well-designed representation will present a high overlap between its map of saliences and its (implicit) map of semantical relevance. Although visual expression is not restricted to pre-attention, the design of visual representations is supposed to maximize design-specific pre-attentive effects. In consequence, there are many graphical properties, but just a limited number of them are used for visual analysis [3]. More incisively, Ware [17] states that understanding what is processed pre-attentively is probably the most important contribution that vision science can bring to data visualization.

The features that are potentially pre-attentive span to a large set. We present an ideally exhaustive set of these features in Table 1. The classification we introduce is reductionist, nevertheless, the literature presents evidence that our generalization is coherent. About *color and texture*, Watt [18] affirms that, just like texture, color is the

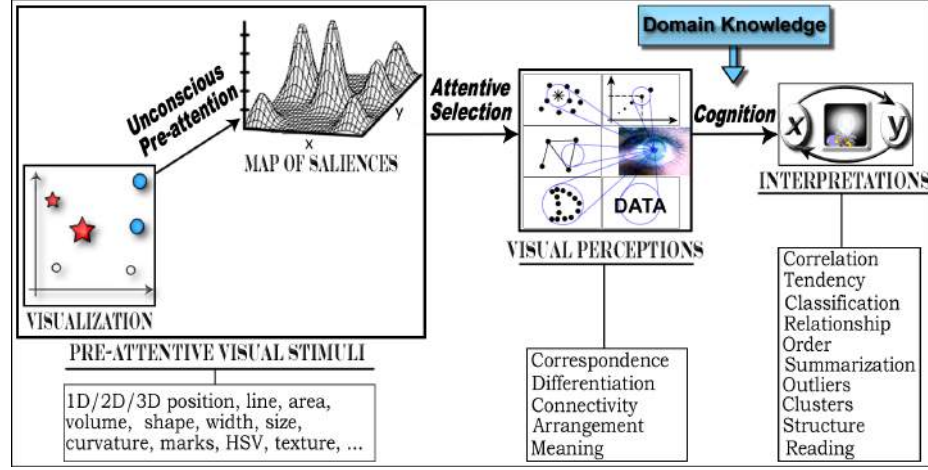


Fig. 3: The Visual Expression Process for visual data analysis.

psychological response to the spectral characteristics of a surface; and that, different surfaces are perceived as having different colors. This is specially verifiable for visual data representations. Motter [12] observes that, early in visual processing, the incoming information is sorted and grouped according to the similarity of simple *shape features*, such as *orientation* or *size*, and of surface features, such as *color*, *luminance*, or *texture*.

Table 1: Classes of pre-attentive features from the perspective of data representation.

Pre-attentive classes	Features
Position	1D/2D/3D position, stereoscopic depth
Shape	line, area, volume, form, orientation, length, width, collinearity, size, curvature, marks, numerosity, convex/concave
Color	hue, saturation, brightness and texture
Animation	movement and intermittence

Pre-attentiveness usually refers to physical properties of particular visual stimuli. Meanwhile, *salientness* refers to the attentional results of a scene as a whole. Each conception has its adequacies and they do not cancel each other. Either way, the fact is that when a set of visual features pops up *pre-attentively*, the brain builds a *map of saliencies* of the scene.

“Effective visualizations must be composed of features that are potentially, and desirably, *pre-attentive*.”

3.2 Visual Perceptions

Pre-attentive features provide maps of potential targets; following this stage, the next vision mechanism is attentive selection. Biased by user intention, one in many of the

prominent areas in a visualization will monopolize the neuronal chain of vision. In what ways the targets of attention are perceived in a visual design?

We have tracked the possible ways in which visual manifestation occurs when the goal is data representation. We have extensively inspected the literature and found a limited set of possibilities in disjoint works. Here we put these findings together in a *visual vocabulary* whose elements appear recurrently over visualization techniques. We call these elements *visual perceptions*.

Visual perceptions are the traits that any user attentively seeks for in a visual representation. This notion is evident when they are not found and the pipeline outlined in Figure 3 is broken, preventing visual expressivity. We have observed that each visual perception is fired by one or multiple pre-attentive stimuli, which are discrete or continuous. According to the model we propose, visual perceptions occur after pre-attention and before data interpretation, independently of the data domain. They bridge vision and data examination.

Our investigation indicates that the elements used in the second stage of visual expressivity are *correspondence*, *differentiation*, *connectivity*, *arrangement* and *meaning*. The most explored and verified of these phenomena, *correspondence* and *differentiation*, are noted by Bertin [2] and by Card *et al.* [3]. The third visual perception is presented by Mackinlay [11] who states that the notion of relationship among graphical entities comes from the perception of *connectivity*. Meanwhile, *Arrangement* arises from group positional configurations, largely studied by the Gestalt psychology [9]. The last perception is *meaning*, that is, resemblance to previous knowledge and/or expertise, studied in psychological models. These five abilities seem to be natural to human beings, more specifically:

- *correspondence*: each position/shape/color has a direct correspondence to a referential map that is part of the scene (explicit) or that is mental (implicit). Explicit maps include axes, geographical maps, shape/color dictionaries, and position/shape/color ranges. Implicit maps include known orderings and shape metaphors;
- *differentiation*: each position/shape/color discriminates graphical items. Differentiation is a correspondence in which the user creates a temporary map in memory;
- *connectivity*: shapes, mainly edges, that convey information about relationships;
- *arrangement*: Gestalt principles of organization. Positional placements (closure, proximity and symmetry) that convey perception about group properties, for example clusters and structural cues;
- *meaning*: positions/shapes/colors whose decoding comes from the expertise of the user or from previous knowledge. Meaning is a correspondence established from visual entities to concepts retained in long-term memory.

“Although visualization researchers have worked on the specificity of different designs [14], apparently, techniques of any nature explore means to express the limited vocabulary of *visual perceptions*.”

3.3 Interpretations

The last step of the Visual Expression Process is to generate interpretations, the goal of visual data analysis. Interpretations refer to deductions, inferences or conclusions created when *visual perceptions* are considered in the context of the *data domain*.

The working memory, the long-term memory and the sensorial system support the generation of interpretations. Working memory has a limited capacity with only three to seven positions (chunks) to be occupied simultaneously and to support cognition. It can be filled with data from the long-term memory or from the sensorial system. However, maintaining data from the long-term memory in the working memory, for more than 3 seconds, demands cognitive resources. Differently, sensorial data is maintained in working memory without demanding cognitive resources; nonetheless, it asks for an effort of attention (sustained attentive selection).

The visual sensorial system, in turn, provides spatial information at rates much higher than the manipulation of images previously stored in long-term memory [17]. Besides, the access of the vision system to the working memory is so rapid that it rivals to the transferring of information from the long-term to the working memory. Filling the working memory with data, either from the long-term memory or from the sensorial system, takes between 100 and 250 *ms* [8].

Visual stimuli, assisted by visual perceptions, work similarly to the images stored in long-term memory. If data interpretation directly relates to cognition, then, *computer graphics make us smarter*. Interpretations are organically semantical; a suggestive set of these phenomena includes: correlation, tendency, classification, relationship, order, summarization, outlier, cluster, structure and reading.

“*Visual perceptions*, when in short-term memory, are conjoined with long-term domain knowledge to produce data interpretations. Indeed, successful *interpretations* necessarily emerge from the vocabulary of *visual perceptions*. For this reason, new principles for design could be investigated from a perceptual perspective.”

3.4 The Process of Visual Expressivity

Figure 3 presents the Process of Visual Expressivity, which describes how visual representations can promote data investigation. It starts from pre-attentive stimuli and, with the aid of visual perceptions, it completes as data interpretations are achieved. Pre-attentive features of position, shape and color relate to raw data. Through the vocabulary of visual perceptions, interpretations are achieved. Interpretations relate not to data, but to new information in the context of the application domain. These three facts relate to vision mechanisms of saliences perception, of attentive selection over interesting targets, and of memory-based cognition.

4 Discussion

The Visual Expression Process breaks visual data analysis into its discrete constituents in order to explain why visualizations work from a visual-perceptual perspective. Given the fact that the Visual Expression Process derives directly from concepts of the vision system, we believe it is the underlying course supporting visualization techniques.

Visualizations unexceptionally depend on a small vocabulary of perceptions independently of the data domain. Visual perceptions define a distinct concept that, we believe, can promote a better support for consciousness in analyzing visual representations. This thesis is justified by the properties of visual perceptions, that:

- are common to every visualization, in contrast to interpretations;

- are not numerous, in contrast to the amplitude of visual patterns that manifest through position, shape and color;
- can be categorically related to the pre-attentive stimuli, allowing for immediate recognition and association;
- constitute a key element for visual analytical cognition.

Such comprehension can foment new approaches to the science of visual representations. In [13], we present a taxonomical analysis that classifies visual representation techniques according to how they employ position, shape and color. In another work [14], we concentrate on the process of spatialization as the determinant for data encoding; the work introduces an alternative design space and a general design model based on multiple cycles of spatialization. Complementing these previous papers, in this work we present evidences from the vision science to support the concepts that constitute the Visual Expression Process.

5 Conclusions

We have traced the connection between *vision science* and *visual representations* from the perspective of *data analysis*. Our model, named Visual Expression Process, departs from pre-attentive features and culminates on data interpretations; it has, as its core, a vocabulary of visual perceptions

The text formulates a piece of communication intended to disseminate vision theory among visualization researchers. We teach concepts to furnish thoughts on how to develop principles for visual representations based on the notion of how they work. The goal is to set a perspective of visual data analysis in which smaller concepts transcribe existing problems, reaching solutions and generating new hypotheses.

Aknowledgements Supported by: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

References

1. A.D. Baddeley and G.J. Hitch. Working memory. *The psychology of learning and motivation: advances in research and theory*, 8:47–89, 1974.
2. J. Bertin. *Graphics and Graphic Information-Processing*. Walter de Gruyter, Berlin, 1977/1981 - 273 pages.
3. Stuart Card, Jock Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA, EUA: Morgan Kaufmann, 1999.
4. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Machine Intel.*, 20(11):1254–1259, 1998.
5. Y. Jiang, I. R. Olson, and M. M. Chun. Organization of visual short-term memory. *Experimental Psychology: Learning, Memory, and Cognition*, 26(3):683–702, 2000.
6. C. Johnson, R. Moorhead, T. Munzner, H. Pfister, P. Rheingans, and T.S. Yoo. *NIH/NSF Visualization Research Challenges*. Los Alamitos, CA, EUA: IEEE Press, 2006. 36 pages.
7. M.A. Just and P.A. Carpenter. A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99:122–149, 1992.
8. D.E. Kieras and D.E. Meyer. An overview of the epic architecture for cognition and performance with application to human-computer interaction. *HCI*, 12:391–438, 2007.
9. Kurt Koffka. *Principles of Gestalt psychology*. International library of psychology, philosophy and scientific method. New York : Harcourt, Brace and Company, 1935.
10. S.J. Luck and E.K. Vogel. The capacity of visual working memory for features and conjunctions. *Nature*, 309:279–281, 1997.

11. J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, 1986.
12. Brad C. Motter. Neural correlates of attentive selection for color or luminance in extrastriate area v4. *The Journal of Neuroscience*, 14(4):2178–2189, 1994.
13. J.F. Rodrigues Jr., A.J.M. Traina, M.C.F. Oliveira, and C. Traina Jr. Reviewing data visualization: an analytical taxonomical study. In *International Conference on Information Visualization*, pages 713–720, Londres, UK, 2006. IEEE Press.
14. J.F. Rodrigues Jr., A.J.M. Traina, M.C.F. Oliveira, and C. Traina Jr. The spatial/perceptual design space: a new comprehension for data visualization. *Information Visualization*, 6:261–279, 2008.
15. J.J. Thomas and K.A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Press, 2005.
16. J.K. Tsotsos. The selective tuning model. 2003. In *Visual Attention Mechanisms*, Kluwer Academic/Plenum Publishers, pages 239–250.
17. Colin Ware. *Information Visualization: Perception for design*. San Francisco, CA: Morgan Kaufman, 2004. 486 pages.
18. Roger J. Watt. Some speculations on the role of texture processing in visual perception. pages 59–67. Massachusetts Institute of Technology, Cambridge, MA, USA, 1995.
19. Richard M. Young and Richard L. Lewis. The soar cognitive architecture and human working memory. pages 224–256. Cambridge University Press, 1999.