

Changing the View: towards the theory of visualisation comprehension

Mark M. Shovman, Andrea Szymkowiak, James L. Bown, Kenneth C. Scott-Brown
University of Abertay Dundee, UK
m.shovman@abertay.ac.uk

Abstract

The core problem of the evaluation of Information Visualisation is that the end product of visualisation – the comprehension of the information from the data – is difficult to measure objectively. This paper outlines a description of visualisation comprehension based on two existing theories of perception: Principles of Perceptual Organisation and the Reverse Hierarchy Theory. The resulting account of the processes involved in visualisation comprehension enables evaluation that is not only objective, but also non-comparative, providing an absolute efficiency classification. Finally, as a sample application of this approach, an experiment studying the benefits of interactivity in 3D scatterplots is presented.

1. Introduction

A review of the methods currently used in evaluating visualisation tools and techniques shows they fall into the following major categories: comparison of performance for similar tasks with different visualisations; user evaluations; and case studies [1]. None of these are perfect, mostly because they are highly subjective and interpretation-dependent [1]. Some objective measures exist in comparison studies, but the problem with comparison is that: (a) it becomes prohibitive in time when considering more than a few items; (b) the sole result of such a comparison is an order relation in a set; (c) it could result in a non-linear ordering (A is better than B, B is better than C, but C is better than A) which may be difficult to interpret, and (d) it is often difficult to generalise beyond the specific task.

The core problem, as various researchers have pointed out [1][2][3][4], is that the object of evaluation, i.e. the end product and the desired aim of Information Visualisation are difficult to define. This might be partly due to the qualitative difference between the data and the information: the data is a concrete measurable object, as is the visualisation; information, on the other hand, is created while actively engaging with the visualised data, and is more related to comprehending the meaning of the data. All these – comprehension, meaning, and

information – are mental constructs, not physical properties, and as such are difficult to define and even more difficult to measure objectively¹ – difficult, but not impossible.

For more than a century, Cognitive Psychology has generated, tested and refined models of various mental processes, resulting in rigorous paradigms for describing mental constructs in an objective, measurable manner. Within this framework, accounting for processes involved in Information Visualisation comprehension becomes possible.

It has been noted that “*Graphs are a recent invention, and if they are an especially effective method of communication, it is because they exploit general cognitive and perceptual mechanisms effectively*”² [2]. In that case, existing theories of perception and comprehension may be applied to describe graph comprehension, and any such description may then be verified by established methods developed within these fields. A coherent account of the processes involved in the Information Visualisation comprehension would then facilitate the measurement of comprehension efficiency, which is a key quality criterion of an Information Visualisation.

It might seem at this point that the authors fail to notice the considerable influence Cognitive Psychology already has on Information Visualisation. This is not the case. Findings relating to human perception, e.g. the amount of distinct colours a human being is able to perceive, are commonly taken into account when designing novel visualisations (e.g [3]). Experimental methodologies developed for assessing human perception can inspire experiment design for Information Visualisation evaluation (e.g. [4]). However, using cognitive theories to outline processes of Information Visualisation

¹ A similar problem can be observed in Computer Graphics, where the drive towards photorealism has hit the uncanny valley [5]. Apparently, the real quality criterion of Computer Graphics is not photorealism, but the perceived realism, the amount of inner belief in the reality of a scene.

² This claim has been also made about reading, and has been receiving numerous supporting evidence since then.

comprehension is, to the best of the authors' knowledge, a largely underexploited area.

2. The pop-out of meaning

It is not that the theories of human perception are unknown to Information Visualisation developers. For example, the Laws of Gestalt [6] are widely quoted (e.g. [3], [7]) – but these are almost a century old. Meanwhile recent developments, even if highly relevant, go relatively unnoticed. The Principles of Perceptual Organization (PPO) [8] are not just a set of loosely-connected laws. They are a coherent account of perception as a process of parsing sensory input into increasingly meaningful and abstract objects through iterative processes of grouping, figure-ground segregation and formation of illusory contours [8]. Most importantly, these Principles come with objective criteria of gestalt-ness – a vast improvement over Wertheimer's introspective observation.

Another well-known theory of perception is the Feature Integration Theory (FIT) [9]. It is based upon an observation of what became to be known as the 'pop-out effect'. In some cases, a visual search task for an 'odd-one-out' element in a stimulus array takes longer for larger stimulus arrays; this effect is considered a consequence of a 'serial search' for the target. In other cases, the size of the stimulus array does not affect reaction time at all; this has been termed the 'pop-out effect' due to the subjective perception of the odd-one-out 'popping out' at the observer. According to the FIT, 'pop-out' happens for 'primitive features' such as colour, orientation, intensity, curvature and shape, that are detected in the low-level³ visual areas, and does not happen for combinations of such features, termed 'feature conjunctions', that can only be perceived further up the hierarchy of visuo-perceptual processing.

Since the early 1980's there has been a growing body of evidence incompatible with this theory, showing 'pop-out' for objects differing in high-level semantics, rather than low-level primitive features. It has been shown that faces 'pop out' from buildings or tools [10], (see Figure 1), and that within very similar arrangements of 'primitive features', some exhibit 'pop-out', while others do not (see Figure 2).



Figure 1: pop-out of high-level features

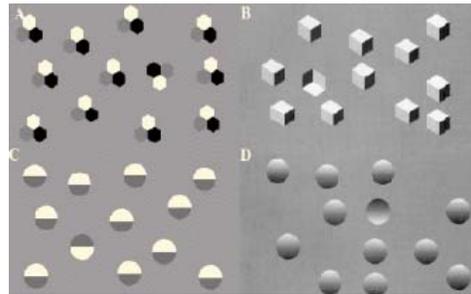


Figure 2: feature conjunction 'pop-out'. left: serial search; right: 'pop-out'

Recently, a theory has been proposed that explains that evidence. Reverse Hierarchy Theory (RHT) [11] dissociates between fast conscious perception and low-level vision. The 'pop-out' effect is attributed to the processes taking place in high areas, where large receptive fields underlie spread attention detecting categorical differences. The search for conjunctions or fine discriminations, on the other hand, is assumed to depend on re-entry to low-level specific receptive fields using mechanisms of serial focused attention [11].

RHT proposes that perception occurs in two distinct stages. First, the 'bottom-up' perception, leading to increasingly complex representations, is automatic and implicit. Then the conscious decision, originating at the hierarchy's top, gradually drives attention downward as needed. Thus, our initial conscious percept matches a high-level, generalized, categorical scene interpretation. For later vision with scrutiny, attention is focused at specific, active, low-level units, incorporating into conscious perception detailed information available there. That two-stage outlook at perception makes RHT highly relevant to the question of Information Visualisation evaluation.

3. Visualisation Comprehension

Both PPO and RHT view perception as a process resulting in the formation of high-level abstract understanding of the sensory stimulus. When that stimulus is a visualisation, it is reasonable to assume that this high-level abstract understanding, ideally, should be the information, i.e., the meaning of the data.

³ In psychology, the sensory organs are often referred to as the 'bottom' of the mind and consciousness and reasoning – as the 'top'

This, together with RHT's distinction between fast implicit 'bottom-up' and slow explicit 'top-down' processes, gives us a distinction between efficient and inefficient visualisation – a quality criterion. An efficient visualisation is one that allows implicit perception of required information; in an inefficient visualisation, the information has to be searched for, explicitly, laboriously and slowly. A good visualisation technique is one in which the message 'pops out'; in a bad one it can only be found after one-by-one scrutiny.

As an added bonus, this criterion comes with its own evaluation methodology that is objective, repeatable and non-comparative. In other words, it is possible to evaluate a visualisation tool or technique that is standalone yet may be compared objectively across independent experimental designs. This could be done by gathering performance indicators, usually response time, in relation to the amount of data presented, and checking whether this relationship is linearly increasing or constant.

4. A Sample Experiment

In the following section, the applicability of this approach is demonstrated by evaluating the benefits of interactivity in 3D visualisations. This particular research question was partly inspired by Colin Ware's observation that "*There has been little or no empirical work on the role of depth cues in perceiving structures such as clusters and correlations in 3D*" [3].

The experiment tested whether a 'pop-out' effect occurs in a common information visualisation task, i.e., outlier detection in a scatterplot, either with or without interactive control over the viewing angle. The apparatus was a standard office desktop PC with an LCD screen and a USB mouse, running a custom-made application written in Java with Java3D.

Participants (11f/9m) were asked to point out an outlier in a cloud of multi-sized multi-coloured spheres. In half the trials, the presented scene was slowly rotating around the vertical axis within a 45° aperture; in the other half, participants could use a mouse to rotate the scene within the same limits themselves.

In most cases, the outlier was defined by its position – 3 standard deviations away from the mean (see Figure 3). In some cases, though, the outlier was defined by a different colour. These colour outlier cases served as the positive control, since a clear 'pop-out' effect was expected consistent with existing literature [12].

The stimulus set size was varied – 27, 64, 125, 216 or 343 spheres, keeping the density constant. Colour outlier stimuli were in three sizes only – 27, 125 and 343 spheres. Each of the pre-generated 64 stimuli (40 with a position outlier and 24 with a colour outlier) was shown twice to each participant, once with interactive control of the rotation, and once without.

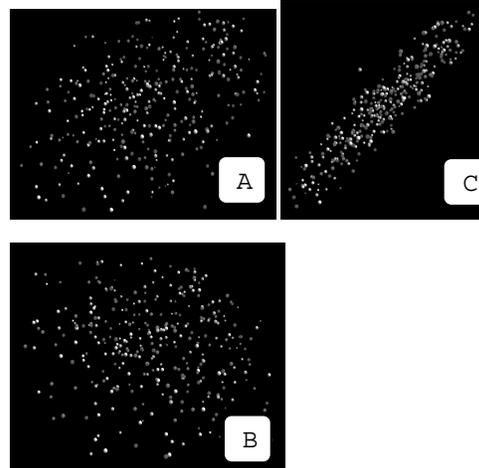


Figure 3: A single 343-sphere stimulus. A&B: left and right limits of rotation; C: side view with the target outlier clearly visible. Note: the side view was never shown to participants

Full logs of mouse movement were analysed to separate the time spent in rotating the target (the interaction time) from the time spent moving towards the target to respond, time spent watching the non-moving scene, etc. Only the trials where the outlier was correctly identified were analysed. The analysis reported on here is the effect of stimulus set size on interaction time (see Figure 4).

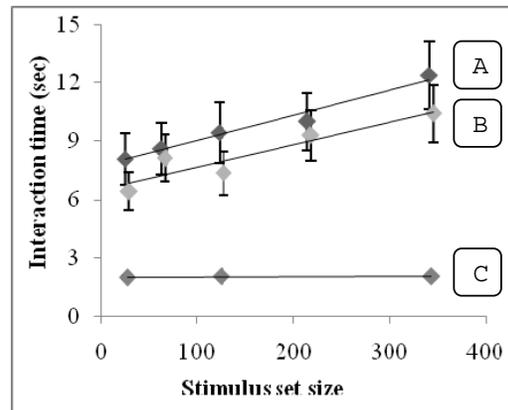


Figure 4: Effect of stimulus set size on interaction time. A: no interactivity; B: interactive rotation; C: colour outlier

For the colour outlier trials (C), the number of spheres had no significant effect on the interaction time (a slant of 0.2msec/item) which indicates a clear 'pop-out' effect. However, for the position outliers a significant effect of the number of spheres on interaction time was found: an increase of 11.2msec/item ($p < 0.01$) for interactive rotation (A) and an increase of 6.6msec/item ($p < 0.01$) for non-interactive rotation (B). A pairwise comparison of performance of the same subject on the same stimulus set in different interaction conditions shows

a significant benefit of the interactive rotation: it takes, on the average, one second less to identify an outlier in this condition (mean interaction time for interactive rotation: 5419msec; mean interaction time for automatic rotation: 6446msec; T-test $p=0.01$)

The experiment demonstrated that interactive control of the 3D scene rotation improves interaction time, but not to the extent that outlier detection is immediate. More importantly, the results demonstrated how the use of the ‘pop-out effect’ paradigm not only allows for comparison among different visualisations, but also provides an absolute efficiency classification for each technique on its own.

5. Conclusion

We have presented an account of Information Visualisation comprehension that is based on two established cognitive models of perception and comprehension – PPO and RHT. From this account, an efficiency criterion can be derived: implicit formation of the relevant information from the visualised data. This criterion enables objective and non-comparative evaluation of any given visualisation technique based on the ‘pop-out’ effect. The applicability of this approach has been demonstrated with an example of assessing the added value of interactivity in 3D scatterplots.

References

- [1] Plaisant C (2004). The challenge of information visualization evaluation, *IEEE Proceedings of Advanced Visual Interfaces* 109-116
- [2] Pinker S (1990). A theory of graph comprehension. *Artificial intelligence and the future of testing* 73-126
- [3] Ware C (2004) *Information Visualization: Perception for Design*. Morgan Kaufmann
- [4] Ware C (2003) Design as Applied Perception. *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*. Morgan Kaufmann
- [5] Mori, Masahiro (1970). Bukimi no tani The uncanny valley (K. F. MacDorman & T. Minato, Trans.). *Energy* 7(4) 33–35. (Originally in Japanese)
- [6] Wertheimer M (1922). Untersuchungen zur Lehre von der Gestalt. *Psychological Research* 1(1) 47-58
- [7] Chen C (2004) *Information Visualization: Beyond the Horizon*. Springer
- [8] Pomerantz JR, Kubovy M (1986). Theoretical approaches to perceptual organization: Simplicity and likelihood principles. *Handbook of perception and human performance* 2 1-46
- [9] Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cognitive Psychology* 12(1) 97-136
- [10] Tovee, MJ (1998) Is Face Processing Special? *Neuron* 21 1239-1242
- [11] Hochstein S, Ahissar M (2002) View from the Top: Hierarchies and Reverse Hierarchies in the Visual System. *Neuron* 36(5) 791-804
- [12] Pomerantz, JR (2006). Colour as a gestalt: Pop out with basic features and with conjunctions. *Visual Cognition* 14(4-8) 619-628