Visual cognitive styles in virtual environments

Constructing and evaluating adequate tasks

Pavel Ugwitz, Vojtěch Juřík, Čeněk Šašinka

Department of Psychology, Faculty of Arts MU / Hume Lab

172577@mail.muni.cz, jurik.vojtech@mail.muni.cz, cenek.sasinka@mail.muni.cz

Abstract

Visual cognitive styles manifest in ways humans implicitly process visual cues from their surroundings. A question arises: how are cognitive styles to be measured? Classic approaches make use of questionnaires or figural tasks. However, the nature of these approaches fails at capturing complexity of real visual compositions. With the onset of (mobile) eye-tracking devices, this limitation can be overcome; the paradigm, however, still revolves around receptive processing of existing visual material. The aim of this paper is to consider constructing 3D virtual scenes, followed by utilizing such scenes in visual cognitive styles tasks – using virtual reality and eye-tracking devices. While constructing virtual environments does indeed offer new possibilities, technological constraints and visual perception limitations do, however, introduce some new limitations to the table. Therefore, offering a methodological grasp of virtual scene construction is relevant.

Introduction

As characterized by some of the pioneers in the field of cognitive differences research (Ausburn & Ausburn, 1978), cognitive style is: *a psychological dimension that represents consistencies in how an individual acquires and processes information*. A visual cognitive style is a further, finer differentiation of the former concept. However, since this metaphorical terrain of cognitive styles research has grown ever so complex over the years (Kozhevnikov, 2007), it is also worth mentioning that in this proposal, we chose to adhere to the definition of visual cognitive styles formulated by Kozhevnikov & colleagues (Kozhevnikov et al., 2005); acc. to said researchers, there are categories of object visualizers and spatial visualizers. Their definition is as follows:

Object visualizers	Spatial visualizers
They prefer to construct vivid, concrete and detailed images of individual objects; they rely primarily on visual-object strategies; they do better on object imagery tasks.	These prefer to schematically represent spatial relations of objects and spatial transformations; they rely primarily on visual-spatial strategies; they do better on spatial imagery tasks.

To measure visual cognitive styles, the original authors have been using the *Object-Spatial Imagery questionnaire* (Blajenkova et al., 2006). While taking this into consideration as a baseline, in our research, we intended to propose computerized ways of measuring visual cognitive styles that make use of eye-tracking within 3D virtual environments.

This submission intends to cover the problematics of constructing 3D scenes salient for eye-tracking stimuli, and to go over scene optimization – so as to ensure smoothness of data collection and future data analysis.

Methods of 3D eye-tracking

By using a virtual reality headset with eye-tracking capability, extracting valid visual cognition data out of a real-time 3D virtual world is entirely possible. There are, however, different possibilities of collecting that eye-tracking data (and the choice of eye-tracking solution influences 3D scene creation, and vice versa). The two key solutions to eye-tracking that are to be considered are:

Heatmap tracking	Object tracking
In 2D eye-tracking applications, this approach collects the X/Y coordinates of the eye; after the data collection phase is done, the most/least frequented gaze areas are plotted onto the 2D plane, by adding a colored overlay on top of the plane. With 3D solutions, this gets more complex, as the third dimension adds the Z axis (X/Y/Z), and a non-static movement through the 3D environment introduces an ever-changing viewport. Measurement accuracy is great; however, both technical implementation and data analysis of this approach are challenging.	Eye-tracking happens on the level of 3D objects. An object in the scene has two states: it is either being fixated by eye-tracking, or not. This allows for simplicity in analysis/implementation; on the other hand, when gazing into details of a single object, there is no further data discrimination. I.e., the pros & cons of this are the opposite of the heatmap approach.

Due to the aforementioned complexities of 3D dynamic heatmap tracking, we opted for the object tracking approach. This means that the elements of the scene shall be broken into individual 3D objects, so that data can be collected on said objects.

Experimental design

In the pilot testing phase, there was no division of participants into experimental groups. The participants, however, were screened on their visual cognitive styles disposition by self-administering the OSIVQ questionnaire. Our intention was to compare these results to the performance and behaviour in a 3D scene we'd construct.

The eye-tracking task was a part of a larger test battery – one that happened in 3D to an extent, being supplemented by traditional pen & paper tasks. As for the 3D portion, we utilized the Unity engine (version 5.4), the Oculus Rift DK2 headset, an SMI eye-tracking upgrade for said headset, along with software solution of our own – to facilitate data collection.

Environmental design

To have something worthy of measuring in a scene, the intended scene must be purposely designed in way of it being saturated in stimuli that are relevant to the theory we are following – i.e. to contain measurable elements that engage both visual cognitive styles, differently, and these differences can be recorded (and, possibly, predicted).

For a predominately object-oriented visualizer, outfitting the scene with objects rich enough in detail is crucial. They spend a substantial amount of time on objects that catch their interest – provided said objects are salient enough. As for spatial visualizers, enriching the scene with relationships is the key. The spatial individual does not spend much time on individual objects; instead, they search for rules that bound the elements of the scene into patterns.

Object cues	Spatial cues
Unique. A common shortcoming of computer-generated 3D scenes that these scenes appear bland. This is so due to the creators of such scenes re-using the same 3D objects over and over again. Therefore, the objects of focus shall be made of an unique 3D models. **Breaking the pattern**. Patterned visual stimuli, such as those found in nature, are not very demanding in terms of visual processing/attention (Taylor et al., 2005). Therefore, an object of focus shall aim for coming up with a distinctive external shape (silhouette) and internal patterns. Let these be (somewhat) non-repetitive/non-conventional, too. **Detailed**. An object-oriented participant is inclined to gaze into the details of individual objects. To facilitate this, let the object have unique features of varying details. As it is with art, bland, visually simplistic areas shall complement detailed, somewhat unexpected parts. This should be contained within an object.	**Lead lines**. Visual art is processed as a flow of lines (3-by-3, 2-by-2 division, 1/2/3 point perspective, horizon line, convergence of lines to an "action center", etc.) The participants tend to follow these (guide)lines, or to proceed onto an area of focus (ideally, any scene should have more than one of these areas of focus, so as to prevent pro-object visualizer interpretation). **Scene boundaries**. These distinctive elements should indicate where an intended scene ends. **Similarities**. If there are multiple objects in the scene with distinctively similar features, these will form a logical group. The spatial visualizer is expected to notice the pattern of similarities, and to follow these groupings of objects. **Encompassion**. This can be achieved by placing a large object that spans across a significant area of the scene, followed by inserting supplementary objects into the body of the large one – all in a way that respects, and complements, the shape and the flow of the large object.

Figure 1: **(left)** An object stimulus (a flag with detailed emblems in it). **(right)** An overall spatial layout of the scene.

Data (pre)analysis

In object-tracking, if we wish to gain extra data from an object in the scene prior to data collection, it has to be broken into multiple sub-objects, while retaining its appearance of a whole object.

To prevent inaccuracies and/or data noise, it is recommended not to overlay eye-tracked 3D objects onto other 3D objects; with the added depth axis of 3D imagery, it could be difficult to determine the object of focus, should edges of multiple objects cross. In fact, eliminating the depth axis to an extent is considered preferable; this can be achieved by placing the objects into the scene, followed by placing a monolithic barrier of constant depth, beyond these eye-tracked objects. Also, to consider:

- **Short distance** of objects. Some 30m may be considered maximum (depending on object size).
- **Object size**. Noticeable and trackable objects should take up at least 5% screen width/height.
- **Proper scene lighting** should complement object recognizability.
- **Objects on the edge** of the scene will be implicitly harder to notice than objects in the middle.
- **Static camera**. To retain control over the experimental design.
- **FOV of the scene**, and the potential need for camera rotation (110 degrees, for most VR headsets).

The aforementioned reduces the potential for data noise. Furthermore, adding tolerance to object boundaries and defining minimum eye-tracking fixation times, data clarity can be achieved.

Conclusions

We have developed a platform that allows for collecting behavioral data in VR; within the scope of virtual environments, we have proposed a way of creating eye-trackable 3D scenes. The process of creation was intended to respect the visual cognitive styles theory. Also, potential technological shortcomings were covered – so that data collection/cleansing are not negatively affected.

We are yet to analyse data and produce results based on the eye-tracking task we have implemented. For this, data processing solutions may have to be developed. The proposed approach, however, serves as a research design baseline for future studies in visual cognitive styles and similar subjects.

References

Ausburn, L. J., & Ausburn, F. B. (1978). Cognitive styles: Some information and implications for instructional design. *Educational Communications & Technology Journal*, 26, pp. 337-354.

Blajenkova, O., Kozhevnikov, M., Motes, M. A. (2006). Object-Spatial Imagery: A New Self-Report Imagery Questionnaire. *Applied Cognitive Psychology*, 20, pp. 239–263.

Kozhevnikov, M. (2007). Cognitive Styles in the Context of Modern Psychology: Toward an Integrated Framework of Cognitive Style. *Psychological Bulletin*, 133 (3), pp. 464–481.

Kozhevnikov, M., Kosslyn. S., Shephard, J.(2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition*, 33 (4), pp. 710-726.

Acknowledgements

This submission was made possible due to the following project: Vliv kartografické vizualizace na úspěšnost řešení praktických a výukových prostorových úloh (MUNI/M/0846/2015).