



Master's Thesis

**Was auch nur die
wenigsten wissen
– bis heute unterdrücktes Wissen
um die Wahre Physik**

Richard Metzler

Draft vom May 3, 2017



Master's Thesis

**Was auch nur die
wenigsten wissen
– bis heute unterdrücktes Wissen
um die Wahre Physik**

Richard Metzler

Aufgabensteller:	Prof. Dr. Dieter Kranzlmüller
Betreuer:	Markus Wiedemann
Abgabetermin:	7. Juli 2077

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

München, den 7. Juli 2077

.....
(*Unterschrift des Kandidaten*)

Abstract

Hier steht eine kurze Zusammenfassung der Arbeit. Sie darf auf gar keinen Fall länger als eine Seite sein, ca. eine drittel bis eine halbe Seite ist optimal.

Contents

1	Introduction	1
2	Related work	3
2.1	Mesh saliency and human perception	3
2.2	Human attention in Virtual Reality	5
	List of Figures	7
	Bibliography	9

1 Introduction

Beginning in the 1950's, virtual reality technology [[Ste92]] has been continuously researched and improved and its professional relevance is becoming ever more present today. There is a plentitude of recent works showing that it bears great potential and positive possible contributions to architecture and construction [[SM14], [LPP15], [SJRT13]], health-care and psychotherapy [[BB14], [MF14], [dRKH⁺14]], engineering and industrial design [[MEC14], [WCAH⁺16]], gaming and home entertainment [[VCSF16], [Zyd05]] and education [[MGC⁺14], [OF⁺15]]. One must also consider that this technology can help gaining new insights and open up new perspectives into greater, more abstract matters of social, environmental and economic manner [[OHH⁺15], [NNVL⁺16]].

With resources such as memory and computing power becoming more and more available at ever-increasing rates, 3D objects and their mesh representations are constantly growing in complexity and size, in terms of shaders, texture maps as well as the sheer number of vertices. Still, many professional applications revolving around interaction with such models require means of displaying them in real-time without significant perceived loss of quality to ensure a smooth and fast workflow. This is where mesh simplification and segmentation plays an important role [[WL10]], [[SG01]], [[ZLSZ12]] etc.

This issue becomes even more pressing in a professional, commercial context where access to state of the art, high-performance graphic processing units or render farms is not a given for everybody. With less computing power available, means of user-oriented, real-time rendering are of vital importance to a fast and unimpeded way of working on 3 dimensional assets.

Little has been done so far concerning research on *mesh saliency* on a vertex level in a virtual reality environment. Whether saliency maps computed via known methods affect user behavior immersed in VR scenes at all, or to which extent has not been investigated in an object manipulation scenario. Furthermore, as of now, the effect on perceived visual quality of saliency-based object simplification as well as user behavior when given the opportunity to declare salient regions themselves is barely touched on at all.

The basic idea of the virtual reality (VR) CAVE [[CNSD93]] is to create an interactive, immersive virtual reality environment for users by setting up multiple projectors and tracking sensors in a room with three to six solid walls that are also suitable be used as a projection surface. The users wear a pair of stereo shutter glasses which are synchronized with the projectors and thus can separate two images for the spectator - one for each eye. The glasses are equipped with some type of tracking system, for example an electromagnetic one, so that their exact position, orientation and tilt can be captured in real-time, allowing the computation of the users perspective in 3D scene at any time. Based on this perspective, the projectors use the walls as projection surfaces and throw live images resembling what the user would see if he/she were physically in the virtual scene onto the walls.

Depending on the number of projectors used, this usually invokes the issue of the user's shadow greatly impairing the immersion. In order to solve this problem, a CAVE must be carefully planned and set up, usually as an entirely separate room within a building. The

1 Introduction

projectors are then installed behind the walls of that room and their projection is cast onto them from the back.

Development of the CAVE model started in 1992 and has since been a continuous effort towards a professional VR solution that is able to eliminate most common problems that other common setups tended to cause back then - and still do - like the need for wearing an uncomfortable VR-headset attached to a cable, limiting the mobility of the user.

2 Related work

Since this work focuses on the impact that being immersed in an interactive virtual reality scene has on human attention when both focusing and performing tasks on 3D objects, this section will be subdivided into two parts. First, publications discussing *mesh saliency*, a commonly used term to describe regional perceived importance of digital representations of real-world (or synthetic) objects as well as 2D images, in the context of human perception will be presented and briefly described. The second section will provide a look into human behavior, based on cognition and outside stimuli in virtual reality environments. Both of these sets of scientific works will provide a solid knowledge of terms and methods commonly used in this field and describe the current state of the art.

2.1 Mesh saliency and human perception

Research on what human perception guides us to focus our attention on when presented a 3D representation of an object was begun just past the year 2000 and has been a continuous effort ever since. One commonly cited publication in this field is Lee *et al.* [[LVJ05]]. Based on low-level human visual attention [[KU87]], it introduces the term *mesh saliency*, a measure for regional based importance of 3D meshes and also presents a way to compute it. This fully automatic process successfully predicts what would be classified by most observers as prominent, visually interesting regions on a mesh, thus allowing mesh operations such as simplification [[CMS98]] and segmentation [[Sha08]] to produce results that are more appealing to the beholder.

The model for computing *mesh saliency* is based on a center-surround comparison of local curvature. It is scale-dependent on a *saliency factor* ε , which is based on the diagonal of the objects bounding box, and is able to identify salient features of a mesh, depending on their surrounding area. Geometrically complex regions, for example a large patch containing lots of bumps of similar size, will be rightfully dismissed as, in most cases, regions that are not interesting from a human perceptual stance.

Taking a closer look at the basic formula through which saliency for any vertex of a mesh can be computed according to Lee *et al.* helps understanding the underlying concept. As a first step, the mean curvature map for a mesh, describing mean local curvature values on a point-level for each of its vertices, needs to be calculated via commonly known approaches such as [Tau95]. The resulting mean curvature map \mathcal{C} defines a mapping from each vertex of a mesh to its mean curvature $\mathcal{C}(v)$. Using a distance measure such as the Euclidean or geodesic method, one can compute the neighborhood $N(v, \sigma)$ of a vertex v which then defines a set of points within a distance σ . The Euclidean approach was used in Lee *et al.* and subsequently in the formula below. Using these definitions, the authors denote the Gaussian-weighted average of the mean curvature by $G(\mathcal{C}(v), \sigma)$ and present the following way of computing it.

$$G(\mathcal{C}(v), \sigma) = \frac{\sum_{x \in N(v, 2\sigma)} \mathcal{C}(x) \exp[-\|x - v\|^2 / (2\sigma^2)]}{\sum_{x \in N(v, 2\sigma)} \exp[-\|x - v\|^2 / (2\sigma^2)]}$$

For computation of the Gaussian-weighted average, a cut-off factor for the filter is assumed at a distance of 2σ , in other words twice the distance that a vertex can have to another vertex to still be considered in its neighborhood. Based on these definitions, the saliency $\mathcal{S}(v)$ of a vertex v is defined as the absolute difference between the Gaussian-weighted averages, as seen in the formula below.

$$\mathcal{S}(v) = |G(\mathcal{C}(v), \sigma) - G(\mathcal{C}(v), 2\sigma)|$$

In order to get more refined results, one can conduct multiple computations of *mesh saliency* with different values for σ . Lee *et al.* use the previously mentioned *saliency factor* ε with $\varepsilon \in [2, 6]$ in their paper, to generate multiple values for σ .

The concept of *mesh saliency* has since been refined, augmented and adapted to serve as a basis for a multitude of specific use-cases and applications. When processing single vertex saliency, Chong *et al.* [[WSZL13]] took into consideration not only the curvature of the region surrounding the vertex, but also the global context of it. In other words, for each vertex to be attributed a value describing its saliency, its *global rarity*, derived from comparing its features to those of every other vertex of the object, is computed. They performed a user study in which they had participants choose one out of two saliency maps for a set of objects, presented in a random order. One map was generated using their approach, the other one with the model presented in Lee *et al.* Participants were asked to pick the one that was a closer representation of what they would have considered interesting regions and features. Since their method got picked in almost 58 per cent of cases, while the results produced by the model presented in Lee *et al.* were favored in about 42 per cent, this can be considered a true improvement of the way *mesh saliency* can be computed.

One approach to improve the method of finding salient elements in 2D images relied on paying extra attention to depth-information in [[CHR13]]. In this work, Ciptadi *et al.* found that better results in terms of automatic identification of objects and surfaces could indeed be achieved this way. Transferring these insights into a 3D context is easy since visually complex models often base on multiple image-maps describing, among other information, depth values on the surfaces of the model. In [PZV11], the authors took a more task-driven approach to contribute to the concept of *saliency*. Gathering color- and depth information about real-world scenes using a Kinect sensor, they extracted semantic cues about surface heights, relative surface orientations and occluded edges. Based on that data, they computed combined saliency maps which allowed them to assign real-world objects to four different categories, enhancing ways a robotic system can interact with them, providing the best possible points where the objects can be grasped and whether they are in reach at all or not (due to occlusion by other objects).

Another recent work aimed at identifying single, distinct elements and objects of 3D models was presented in [Kos03] by Koschan *et al.* The authors propose a segmentation algorithm that utilizes a human perception phenomenon known as the *minima rule* which suggests that contours of negative curvature minima can serve as boundaries of disjunct visual parts or elements. Another detailed comparison between automatically detected points

of interests and what participants in a study actually declared as visually interesting points was drawn by Dutagaci *et al.* in [DCG12].

To verify the practical relevance of identifying salient regions and features on 3D meshes, Howlett *et al.* [[HHO05]] conducted a user study on whether it is possible to determine such features in advance. Based on observations gathered from eye-tracking device based user studies, they concluded that, especially with natural objects (animals, humans etc.), this was indeed the case. On top of that, they also reported a significant increase in visual fidelity on objects which were simplified based on saliency weight-maps, according to reports of study participants. Furthermore, in [KG03] the authors used user-guided simplification to preserve higher levels of detail in areas of 3D objects that people deemed important to the recognizability of the object. After performing mesh simplification according to [GH97], enhanced by taking user-derived weight maps into account, the authors observed what they described as perceptually improved approximations of input objects.

In another highly noteworthy work [[Mun07]], the concept of *mesh saliency* was extended to deformable, in other words animated, objects. The author presented a way to generate so-called *multi-pose saliency*, a combination of multiple saliency maps computed for static poses of a mesh. These static meshes can also be interpreted as keyframed poses for dynamic deformation, which makes this work a potentially outstanding contribution to any field where 3D objects are being animated.

The author found a remarkable improvement of the original way of computing *mesh saliency* as presented by Lee *et al.* by using geodesic distance [[SSK⁺05]] instead of euclidean distance when comparing local curvature values and implemented a way to compute multiple saliency maps for different levels of detail.

2.2 Human attention in Virtual Reality

In 1998, Atchley *et al.* [[AK]] conducted four experiments addressing attention in 3D scenes. Participants were shown simple scenes, each containing sets of six short lines. The scenes were arranged on four different depth planes, one behind the other, and displayed on a stereographic display. The basic task given to participants in all of the experiments was to focus a briefly visible color singleton on a specific, previously cued depth plane. One of the lines on the indicated plane would change its color for 100 milliseconds and participants had to correctly say whether it was tilted to the left or the right. To determine the time it takes to shift attention from one plane to another, for some user groups, the color singleton would appear in a plane other than the previously hinted one. To further track speed and accuracy of attention focus, a distraction element (one additional line changing its color simultaneously to the target line) was shown to some participants, sometimes in the *target plane*, sometimes in a different one. From their observations, the authors gathered that depth-plane attention can be successfully guided and that distraction elements appearing on the *target plane* significantly interfered with the users' ability to give correct answers, while such elements appearing on other planes had virtually no impact on results.

Taking the effort of tracking attention one step further into virtual reality context, Lee *et al.* [[LKC07]] accomplished just that on an object basis. They presented a framework capable of such a task, both bottom-up (stimulus-driven) and top-down (goal-directed). Based on pixel-level saliency maps, computed via known methods, similar maps for multiple objects in the scene are generated, allowing predictions on which objects will more likely to be focused

first by users. Using the method presented in this paper, object-level saliency maps can be computed in real-time, depending on the users dynamic position and orientation within the scene. The authors exploited knowledge of human cognition which suggests that attention is object-based [[ODK99]] and, using a monocular eye tracking device, compared the results of estimated object-level saliency maps to the behavior of participants in their study. They dynamically assigned saliency values to each object and, depending on how many of the objects with the highest values (first 1, 2 or 3) were taken into consideration, observed estimation accuracies ranging from about 50 per cent to up to nearly 95 per cent. As one can imagine, accuracy values were the highest when users were given a task, for example finding a certain object within the scene. This shows that attention in virtual space can be tracked and accurately predicted on object level.

In [SP08], the authors Sokolov *et al.* took a more general approach to navigation in virtual 3D scenes, aimed at automatic generation of a set of viewports that allow the observer to gain maximum knowledge of the scene. This work provides detailed information on how viewport quality can be determined based upon the relation of the total numbers of visible, occluded and degenerated faces. Sandor *et al.* took the concept of saliency to a augmented reality context in [SCDM10] and presented the second iteration of their method of combining a real-time camera feed with a 3D overlay. Based upon clues from human perception, for each frame, the authors identified visually important objects and prevented them from being occluded or obscured by the 3D scene projected into it.

List of Figures

Bibliography

- [AK] Paul Atchley and Arthur F Kramer. Attentional control within 3-d space.
- [BB14] Oliver Baus and Stéphane Bouchard. Moving from virtual reality exposure-based therapy to augmented reality exposure-based therapy: a review. *Frontiers in human neuroscience*, 8:112, 2014.
- [CHR13] Arridhana Ciptadi, Tucker Hermans, and James M Rehg. An in depth view of saliency. Georgia Institute of Technology, 2013.
- [CMS98] Paolo Cignoni, Claudio Montani, and Roberto Scopigno. A comparison of mesh simplification algorithms. *Computers & Graphics*, 22(1):37–54, 1998.
- [CNSD93] Carolina Cruz-Neira, Daniel J Sandin, and Thomas A DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the cave. 1993.
- [DCG12] Helin Dutagaci, Chun Pan Cheung, and Afzal Godil. Evaluation of 3d interest point detection techniques via human-generated ground truth. *The Visual Computer*, 28(9):901–917, 2012.
- [dRKH⁺14] Sandrine de Ribaupierre, Bill Kapralos, Faizal Haji, Eleni Stroulia, Adam Dubrowski, and Roy Eagleson. Healthcare training enhancement through virtual reality and serious games. In *Virtual, Augmented Reality and Serious Games for Healthcare 1*, pages 9–27. Springer, 2014.
- [GH97] Michael Garland and Paul S Heckbert. Qslim simplification software, 1997.
- [HHO05] Sarah Howlett, John Hamill, and Carol O’Sullivan. Predicting and evaluating saliency for simplified polygonal models. *ACM Transactions on Applied Perception (TAP)*, 2(3):286–308, 2005.
- [KG03] Youngihn Kho and Michael Garland. User-guided simplification. In *Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 123–126. ACM, 2003.
- [Kos03] AF Koschan. Perception-based 3d triangle mesh segmentation using fast marching watersheds. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–II. IEEE, 2003.
- [KU87] Christof Koch and Shimon Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. In *Matters of intelligence*, pages 115–141. Springer, 1987.

- [LKC07] Sungkil Lee, Gerard Jounghyun Kim, and Seungmoon Choi. Real-time tracking of visually attended objects in interactive virtual environments. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 29–38. ACM, 2007.
- [LPP15] Quang Tuan Le, Akeem Pedro, and Chan Sik Park. A social virtual reality based construction safety education system for experiential learning. *Journal of Intelligent & Robotic Systems*, 79(3-4):487, 2015.
- [LVJ05] Chang Ha Lee, Amitabh Varshney, and David W Jacobs. Mesh saliency. In *ACM transactions on graphics (TOG)*, volume 24, pages 659–666. ACM, 2005.
- [MEC14] Stefan Marks, Javier E Estevez, and Andy M Connor. Towards the holodeck: fully immersive virtual reality visualisation of scientific and engineering data. In *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand*, pages 42–47. ACM, 2014.
- [MF14] Alma S Merians and Gerard G Fluet. Rehabilitation applications using virtual reality for persons with residual impairments following stroke. In *Virtual Reality for Physical and Motor Rehabilitation*, pages 119–144. Springer, 2014.
- [MGC⁺14] Zahira Merchant, Ernest T Goetz, Lauren Cifuentes, Wendy Keeney-Kennicutt, and Trina J Davis. Effectiveness of virtual reality-based instruction on students’ learning outcomes in k-12 and higher education: A meta-analysis. *Computers & Education*, 70:29–40, 2014.
- [Mun07] Rodrigo Barni Munaretti. Perceptual guidance in mesh processing and rendering using mesh saliency. 2007.
- [NNVL⁺16] Minh-Tu Nguyen, Hai-Khanh Nguyen, Khanh-Duy Vo-Lam, Xuan-Gieng Nguyen, and Minh-Triet Tran. Applying virtual reality in city planning. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 724–735. Springer, 2016.
- [ODK99] Kathleen M O’craven, Paul E Downing, and Nancy Kanwisher. fmri evidence for objects as the units of attentional selection. *Nature*, 401(6753):584–587, 1999.
- [OF⁺15] Michela Ott, Laura FREINA, et al. A literature review on immersive virtual reality in education: state of the art and perspectives. In *Conference proceedings of eLearning and Software for Education(eLSE)*, number 01, pages 133–141. Universitatea Nationala de Aparare Carol I, 2015.
- [OHH⁺15] Jivka Ovtcharova, Polina Häfner, Victor Häfner, Jurica Katicic, and Christina Vinke. Innovation braucht resourceful humans aufbruch in eine neue arbeit-skultur durch virtual engineering. In *Zukunft der Arbeit in Industrie 4.0*, pages 111–124. Springer, 2015.
- [PZV11] Ekaterina Potapova, Michael Zillich, and Markus Vincze. Learning what matters: combining probabilistic models of 2d and 3d saliency cues. In *International Conference on Computer Vision Systems*, pages 132–142. Springer, 2011.

- [SCDM10] Christian Sandor, Andrew Cunningham, Arindam Dey, and Ville-Veikko Mattila. An augmented reality x-ray system based on visual saliency. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, pages 27–36. IEEE, 2010.
- [SG01] Eric Shaffer and Michael Garland. Efficient adaptive simplification of massive meshes. In *Proceedings of the conference on Visualization'01*, pages 127–134. IEEE Computer Society, 2001.
- [Sha08] Ariel Shamir. A survey on mesh segmentation techniques. In *Computer graphics forum*, volume 27, pages 1539–1556. Wiley Online Library, 2008.
- [SJRT13] R Stouffs, P Janssen, S Roudavski, and B Tunçer. What is happening to virtual and augmented reality applied to architecture? In *Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2013)*, volume 1, page 10, 2013.
- [SM14] Alcínia Z Sampaio and Octávio P Martins. The application of virtual reality technology in the construction of bridge: The cantilever and incremental launching methods. *Automation in construction*, 37:58–67, 2014.
- [SP08] Dmitry Sokolov and Dimitri Plemenos. Virtual world explorations by using topological and semantic knowledge. *The Visual Computer*, 24(3):173–185, 2008.
- [SSK⁺05] Vitaly Surazhsky, Tatiana Surazhsky, Danil Kirsanov, Steven J Gortler, and Hugues Hoppe. Fast exact and approximate geodesics on meshes. In *ACM transactions on graphics (TOG)*, volume 24, pages 553–560. Acm, 2005.
- [Ste92] Jonathan Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, 1992.
- [Tau95] Gabriel Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 902–907. IEEE, 1995.
- [VCSF16] Luis Valente, Esteban Clua, Alexandre Ribeiro Silva, and Bruno Feijó. Live-action virtual reality games. *arXiv preprint arXiv:1601.01645*, 2016.
- [WCAH⁺16] Robert E Wendrich, Kris-Howard Chambers, Wade Al-Halabi, Eric J Seibel, Olaf Grevenstuck, David Ullman, and Hunter G Hoffman. Hybrid design tools in a social virtual reality using networked oculus rift: a feasibility study in remote real-time interaction. In *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages V01BT02A042–V01BT02A042. American Society of Mechanical Engineers, 2016.
- [WL10] Jin Wei and Yu Lou. Feature preserving mesh simplification using feature sensitive metric. *Journal of Computer Science and Technology*, 25(3):595–605, 2010.

Bibliography

- [WSZL13] Jinliang Wu, Xiaoyong Shen, Wei Zhu, and Ligang Liu. Mesh saliency with global rarity. *Graphical Models*, 75(5):255–264, 2013.
- [ZLSZ12] Yitian Zhao, Yonghuai Liu, Ran Song, and Min Zhang. A saliency detection based method for 3d surface simplification. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 889–892. IEEE, 2012.
- [Zyd05] Michael Zyda. From visual simulation to virtual reality to games. *Computer*, 38(9):25–32, 2005.