# Pose and Visual Attention: Exploring the Effects of 3D Shape Near-Isometric Deformations on Gaze

Valeria Garro
Blekinge Institute of Technology
371 79 Karlskrona, Sweden
valeria.garro@bth.se

Veronica Sundstedt
Blekinge Institute of Technology
371 79 Karlskrona, Sweden
veronica.sundstedt@bth.se

## ABSTRACT

Recent research in 3D shape analysis focuses on the study of visual attention on rendered 3D shapes investigating the impact of different factors such as material, illumination, and camera movements. In this paper, we analyze how the pose of a deformable shape affects visual attention. We describe an eye-tracking experiment that studied the influence of different poses of non-rigid 3D shapes on visual attention. The subjects free-viewed a set of 3D shapes rendered in different poses and from different camera views. The fixation maps obtained by the aggregated gaze data were projected onto the 3D shapes and compared at vertex level. The results indicate an impact of the pose for some of the tested shapes and also that view variation influences visual attention. The qualitative analysis of the 3D fixation maps shows high visual focus on the facial regions regardless of the pose, coherent with previous works. The visual attention variation between poses appears to correspond to geometric salient features and semantically salient parts linked to the action represented by the pose.

**Keywords**
Shape Analysis, Mesh Models, Visual Attention, Perception, Gaze Analysis.

## 1 INTRODUCTION

Understanding and modeling human visual attention is a relevant research topic that has been widely investigated in computer graphics and computer vision [Bor13]. Several applications relying on user's gaze detection and analysis have been proposed in the last decades. Foveated or selective rendering [Pat16] and mesh simplification [Lar11] are some examples of computer graphics applications aiming at high perceived visual quality of the scene. In computer vision, many object recognition and object detection applications are based on saliency models [Bor13].

In 3D shape analysis, the concept of *mesh saliency* was introduced in [Lee05]. The authors presented a method to compute a per-vertex saliency measure on the 3D mesh inspired by low-level visual attention. Several other works, e.g. [Wu13, Lei16], proposed similar surface-based saliency models focusing on the geometric properties of the 3D mesh. Recent works provided a step forward into the analysis of visual attention applied to static 3D shapes, investigating the influence of other factors such as different camera views [McD09], rendering lighting conditions, materials, and camera movements [Lav18]. In general, the datasets used in these works consist of rigid 3D shapes such as 3D models of statues, vases, mechanical objects, and non-rigid shapes in resting positions like a quadruped animal standing on four legs. However, in real application scenarios, it is common to deal with deformable shapes that move and assume different poses, in addition to views. For this reason, we focus on investigating visual attention variations when looking at the same shape holding different poses. Our contribution with this paper is to perform a first step to explore how pose could potentially influence visual attention. This is done by acquiring and analyzing eye-tracking data from observers viewing near-isometric deformations of 3D shapes, i.e. human or animal shapes, in different poses.

In the experiment presented in this paper, we asked the participants to free-view a computer screen displaying a set of 3D shapes rendered in different poses and from different angles. We acquired the participants' gaze data using an eye tracker and we computed the 2D aggregated fixation maps. Each 2D fixation map has then been projected onto the related 3D shape transferring the fixation data to the vertices of the mesh and creating a *3D fixation map*. The 3D fixation maps of the same shape in different poses have then been compared to each other to analyze variations in visual attention.

https://doi.org/10.24132/CSRN.2020.3001.18

The remainder of the paper is organized as follows: Section 2 presents previous works on visual attention related to 3D models by eye-tracking analysis; Sections 3 and 4 describe the process to create the stimuli used during the experiment, the equipment used, and the experiment methodology. Section 5 presents the process of creating the 3D fixation maps, while the comparison between 3D fixation maps is described in Section 6. Since we aggregate the fixation data from our valid set of participants, an analysis of the consistency across participants' data (Inter Observer Congruency) is required and it is described in Section 7. The results are analyzed in Section 8, while in Section 9 we summarize the findings of our work.

## 2 RELATED WORK

In the last decades, several works investigated visual attention on 3D objects through the analysis of eye-tracking data focusing on different perspectives. In [How05], the authors, through a set of experiments differing in tasks, analyzed gaze data to determine salient features of 3D models in the context of mesh simplification to maintain high perceptual quality. The sets of 3D models used in these experiments include natural objects, i.e. animal shapes, and artifacts. One of the results showed that the heads of natural objects obtained high values of saliency. Similar findings were presented in [McD09]; this work focused on the analysis of salient body parts of virtual human characters in crowd rendering, revealing high saliency for the heads and the upper torsos. The models displayed in these experiments were textured human characters wearing different casual outfits, either standing in a neutral position or performing walk cycles.

In [Kim10], the mesh saliency model presented by Lee et al. [Lee05] was compared with fixation data acquired through an experiment to verify their similarity. In this experiment, Lee et al. saliency model showed higher correlation values than a random model indicating a correspondence between the saliency model prediction and human eye fixations. Here, the comparison analysis was performed at image level applying a modified version of the chance-adjusted saliency metric [Par02]. More recent works studied visual attention directly on the 3D shapes, both real 3D printed [Wan16, Wan18] and virtual [Lav18]. Saliency models like [Lee05] were also analyzed in [Wan16, Lav18], however in these studies they exhibited poor performance in predicting human fixations, demonstrating the complexity of the visual attention process.

These existing saliency models are based only on the analysis of the geometric features of the shapes. The authors of [Lav18, Wan18] studied also the impact of other conditions on visual attention, such as view orientation and material. In [Lav18], it was reported a significant influence of material, light setting, and camera paths on fixation data when inspecting virtual 3D models. In a setting with real objects [Wan18], different camera views provided different results on fixations; in contrast, the two analyzed materials did not determine any significant difference. Based on these recent findings, we were inspired to go a step further and analyze a property that, to the best of our knowledge, has not studied before, the impact of different poses of non-rigid shapes from different views.

## 3 CREATION OF THE STIMULI

Six non-rigid characters have been selected from the TOSCA high-resolution dataset [Tos, Bro08] for non-rigid shape similarity and correspondence experiments. Meshes of the same character in different poses have the same triangulation and the vertices are listed in the same order. This allows a direct comparison at a per-vertex level. The selected characters are two human males, a horse, a cat, a gorilla, and a centaur. The average number of vertices is about 35000. For each of the chosen characters, four different poses have been selected for a total of 24 different meshes. As shown in Figure 5 and Figure 6, the poses include resting positions (e.g. *Horse0*, *Michael15*), as well as extreme actions (e.g. *Horse10*, *Centaur3*), common (e.g. *Cat2*) and more uncommon (e.g. *Gorilla8*) positions.

Each mesh was rendered from three different camera locations: one in front of the mesh ($V_0$), one at a 45° angle ($V_{45}$), and the last one at 90° angle ($V_{90}$). For meshes belonging to the same character, the cameras were positioned at a constant distance from the center of mass of each mesh. The height of the camera center was set at the same height as the center of mass of the mesh. All meshes were rendered using Blender 2.80[1] Eevee engine, with a Lambertian gray shading. The same lighting setup was used for each camera view, with a single light providing uniform illumination positioned behind the camera and pointing at the same direction of the camera. The final rendered images have a resolution of $1920 \times 1080$.

## 4 EXPERIMENT EQUIPMENT AND PROCEDURE

The stimuli were visualized on a 15.6 inches laptop screen. A Tobii X2-30 eye tracker was used to collect the gaze data of the participants. This device has a sampling rate of 30 Hz and a gaze accuracy of 0.4° under ideal conditions. Tobii Pro Studio software[2] was used to design the experiment and record the gaze data. 21 participants were recruited for the experiment, having normal or corrected-to-normal vision and they were

[1] https://www.blender.org

[2] https://www.tobiipro.com/product-listing/tobii-pro-studio/

not aware of the purpose of the experiment. Participants' age ranged from 19 to 46. The reported genders were 7 F, 13 M, and 1 "Prefer not to answer". The data collected from four participants were rejected due to calibration issues or a low percentage of valid detected gaze. After reading the experiment instructions and signing the consent form, the session started with a 9-points calibration procedure for the eye-tracker device. During the 9-points calibration, the participant was asked to follow a moving red dot on the screen, this procedure lasts for around 40 seconds. As an additional check of the calibration accuracy, the first stimuli set were four images of a cross positioned in different areas of the screen. Each image was shown for four seconds and the participants were asked to look at the center of each cross. Thereafter, the participants were asked to free-view the set of 72 rendered images, each for seven seconds with a monochromatic (mid-gray) image in between the stimuli of the duration of two seconds to separate stimuli reactions. The order in which the rendered images were displayed was randomized for each participant to prevent an order effect bias [Cun11].

## 5 3D FIXATION MAPS

The first step for the creation of the 3D fixation maps is collecting the aggregated gaze data for each stimulus. The identification of fixations has been performed applying the I-VT fixation filter provided by Tobii which classifies eye movements based on their velocity [Kom10]. For each mesh $C_i$, representing the character $C$ in pose $i$, and each view $j$, the set of fixations of all participants were aggregated in a 2D fixation map $f_{ij}^C$, a 2D matrix of the same size of the stimuli. Each fixation contributes by adding to $f_{ij}^C$ a two-dimensional gaussian kernel centered on the pixel coordinates of the fixation point with a radius of 62 pixels which corresponds in our setup to a visual angle of $1^\circ$, i.e. the radius of the fovea, the region in the visual field with highest visual acuity [Duc17]. The values of the 2D fixation map are mapped to the unit interval creating a grayscale image.

We project the 2D fixation image on the related 3D mesh with Meshlab software [Mes]. The current 3D mesh has now grayscale color values at vertex level representing the fixation values, with bright color values corresponding to areas related to high fixation values and dark color values indicating low or absent fixations. The 3D fixation map of the character $C$ in pose $i$ from view $j$ is defined as the list of vertex color values. Since the shapes belong to the TOSCA high-resolution dataset, 3D meshes of the same character have the vertices listed in the same order. Hence, 3D fixation maps of the same character can be directly compared with a chosen similarity or distance metric.

## 6 3D FIXATION MAPS COMPARISON

Since the goal is to analyze the effects on visual attention of different poses of the same character, we compared all pairs of 3D fixation maps of the same character $C$ obtained from the same view $j$. Pearson's Correlation Coefficient (PCC) is used as a comparison measure. For two 3D fixation maps $X$ and $Y$, PCC is defined as $\text{cov}(X,Y)/(\sigma_X \sigma_Y)$, where cov indicates the covariance and $\sigma$ the standard deviation. PCC estimates the linear relationship between two 3D fixation maps and its values range between $-1$ and 1, with 0 implying no correlation and 1 and $-1$ perfect positive and negative linear relationship respectively. PCC is an evaluation metric commonly applied to compute the similarity between 2D saliency maps [Byl19] and it has also been used in recent works on fixation maps applied to 3D shapes [Lav18].

We define $V_{aj}^C$ the set of vertices of the mesh $C_a$ visible from view $j$. When comparing the 3D fixation maps $F_{aj}^C$ and $F_{bj}^C$ related to meshes $C_a$ and $C_b$, PCC is computed exclusively on $V_{aj}^C \cap V_{bj}^C$, i.e. the subset of vertices that are visible from view $j$ for both meshes $C_a$ and $C_b$, to assure the comparison is run solely on valid fixation values.

Due to the variability of the poses, it is unlikely that the observer will look at the exact same portion of the mesh when viewing two different poses of the same character. For this reason, in addition to PCC, the Jaccard Index $J$ (intersection over union) of the two sets of visible vertices $V_{aj}^C$ and $V_{bj}^C$ is computed to measure their similarity: $J = |V_{aj}^C \cap V_{bj}^C| / |V_{aj}^C \cup V_{bj}^C|$. $J$ values range from 0 (empty intersection) to 1 (equal sets). In this context, since PCC is computed on the intersection set, $J$ indicates a measure of the extent of this common region on which the corresponding PCC was measured in relation to the union set of visible vertices. $J$ indicates also how much the two poses vary from each other. If J is close to 0, it means that the two poses share a small set of vertices visible from the same view, due for example to auto-occlusions (e.g. the face of *Gorilla8* is hidden by the left arm from view $V_{90}$) or different nature of the poses (e.g. from view $V_0$, the abdomen of the cat character is visible in *Cat1* but completely hidden in *Cat0*), hence the observer looks overall at different portions of the same mesh. While if J is close to 1, it means that the sets of visible vertices of the two poses are almost congruent, hence the observer looks overall at a similar portion of the mesh.

## 7 INTER OBSERVER CONGRUENCY ANALYSIS

The similarity values obtained from the 3D fixation maps comparison rely on the aggregated gaze data gathered from all valid participants ($N = 17$). To investigate

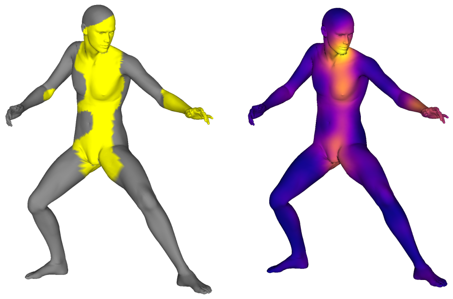


Figure 1: On the left, an example of binary fixation map with threshold 0.2 for *David10* obtained from view $V_{45}$. On the right, the corresponding original 3D fixation map color coded for visualization purpose, with yellow representing the highest number of fixations and blue indicating areas with no fixations.

the *Inter Observer Congruency* (IOC), i.e. the consistency across participants data, we adopt a *leave one out* approach similar to [Tor06]. For each stimulus, the congruency of the 3D fixation map generated from the data of one subject is tested against the partial aggregated 3D fixation map obtained by all the other $N-1$ participants. The final IOC value is generated by averaging the congruency values obtained with this procedure from all subjects.

A binary map indicating the most fixated vertices is created by setting a threshold to the partial aggregated 3D fixation map of $N-1$ participants. The vertices with fixation value $> 0.2$ have been selected creating the binary map which covers most of the aggregated fixations, as shown in Figure 1.

The fixations obtained by the left-out subject are then projected on the 3D mesh as circular patches to take into account eye tracker accuracy errors. The congruency value of the left-out subject is computed as the ratio of the vertices touched by a single observer fixations that fall also within the binary map.

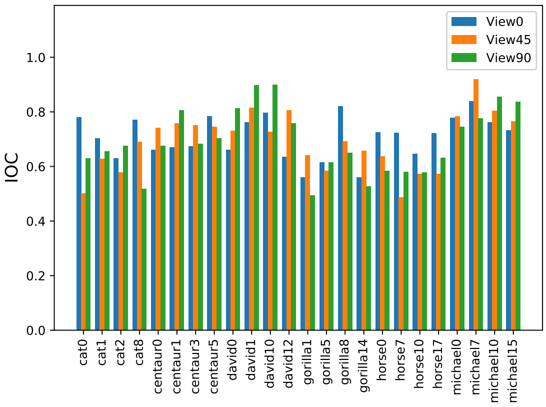


Figure 2: Inter Observer Congruency (IOC) mean values of all stimuli.

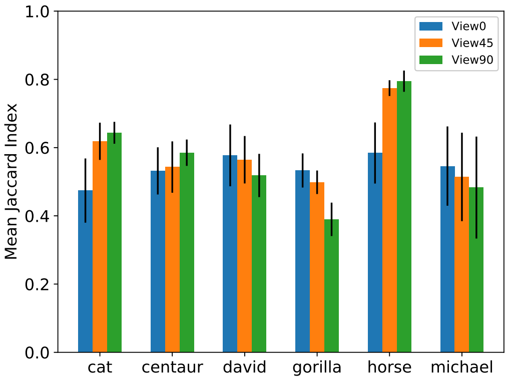


Figure 3: Mean and standard deviation values of Jaccard Index between pairs of shapes grouped by characters and views.

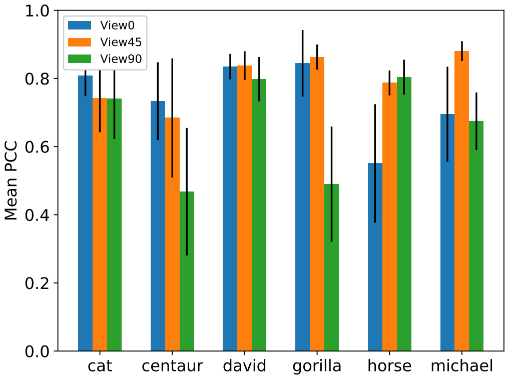


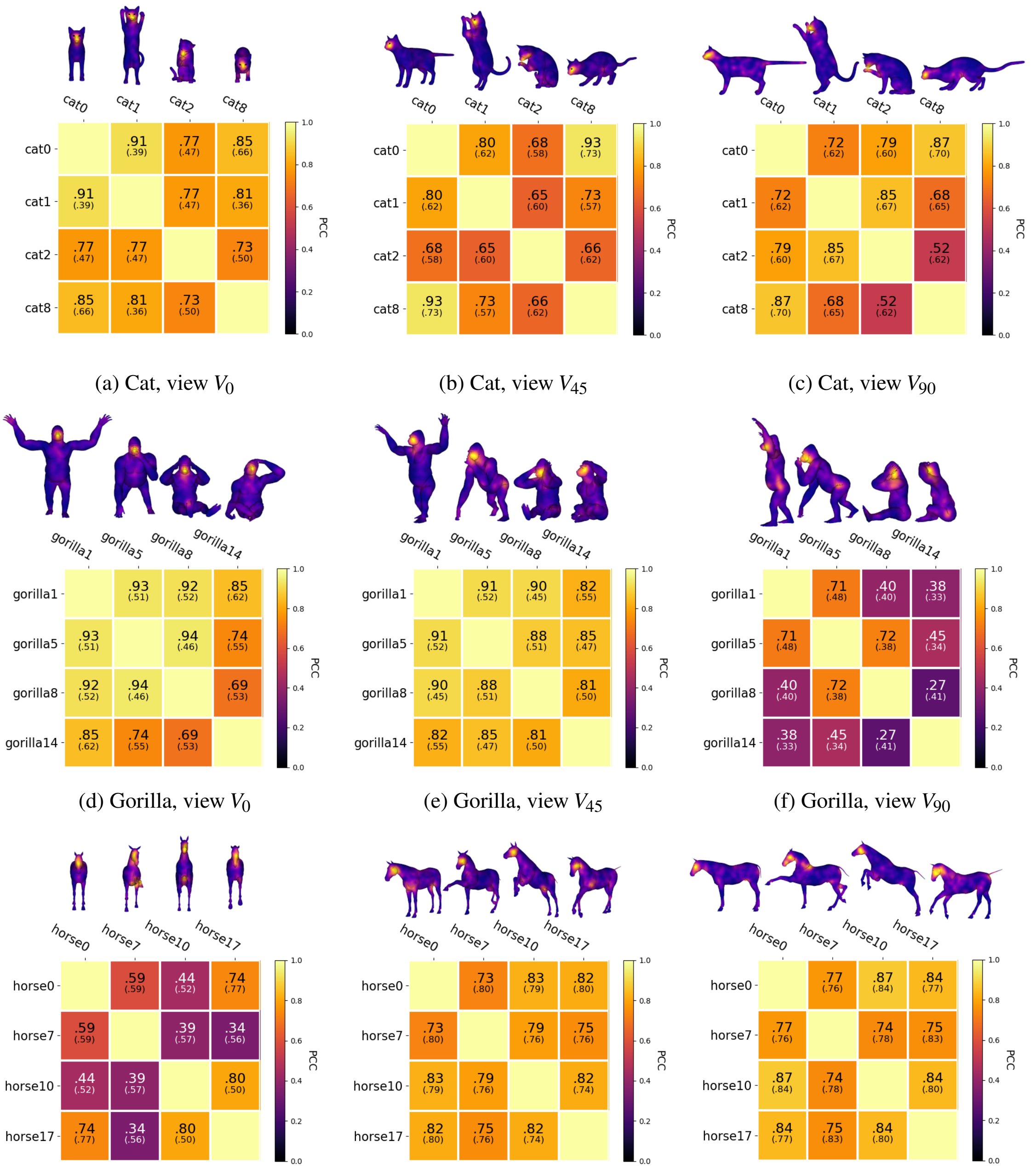
Figure 4: Mean and standard deviation values of Pearson's correlation coefficient grouped by characters and views.

## 8 RESULTS

The IOC results show a total mean value of 0.70 characterizing a fairly coherent dataset of fixations. Hence, the 3D fixation maps computed from aggregating the fixations data of all valid participants can be used on the comparison analysis.

A further analysis of the IOC values show a small variation between views ($\mathrm{IOC}_{V_0} = 0.71$, $\mathrm{IOC}_{V_{45}} = 0.69$, and $\mathrm{IOC}_{V_{90}} = 0.69$) and a wider variation is presented when computing the mean IOC values of the different characters regardless of the view ($\mathrm{IOC}_{\mathrm{gorilla}} = 0.62$, $\mathrm{IOC}_{\mathrm{horse}} = 0.62$, $\mathrm{IOC}_{\mathrm{cat}} = 0.65$, $\mathrm{IOC}_{\mathrm{centaur}} = 0.71$, $\mathrm{IOC}_{\mathrm{david}} = 0.78$, $\mathrm{IOC}_{\mathrm{michael}} = 0.80$). Interestingly, the highest IOC values are the ones related to human characters showing higher consistency across participants when looking at humans. The mean IOC value for each stimulus is shown in Figure 2.

A first qualitative analysis of the 3D fixation maps supports previous findings [How05, McD09] of high visual attention values, across all three views, over characters' heads and human characters' torsos, as shown in Figure 5 and Figure 6. This seems to happen also if changing the pose.



|  | $cat_0$ | $cat_1$ | $cat_2$ | $cat_8$ |
|---|---|---|---|---|
| cat0 |  | .91 (.39) | .77 (.47) | .85 (.66) |
| cat1 | .91 (.39) |  | .77 (.47) | .81 (.36) |
| cat2 | .77 (.47) | .77 (.47) |  | .73 (.50) |
| cat8 | .85 (.66) | .81 (.36) | .73 (.50) |  |

(a) Cat, view $V_0$

|  | $cat_0$ | $cat_1$ | $cat_2$ | $cat_8$ |
|---|---|---|---|---|
| cat0 |  | .80 (.62) | .68 (.58) | .93 (.73) |
| cat1 | .80 (.62) |  | .65 (.60) | .73 (.57) |
| cat2 | .68 (.58) | .65 (.60) |  | .66 (.62) |
| cat8 | .93 (.73) | .73 (.57) | .66 (.62) |  |

(b) Cat, view $V_{45}$

|  | $cat_0$ | $cat_1$ | $cat_2$ | $cat_8$ |
|---|---|---|---|---|
| cat0 |  | .72 (.62) | .79 (.60) | .87 (.70) |
| cat1 | .72 (.62) |  | .85 (.67) | .68 (.65) |
| cat2 | .79 (.60) | .85 (.67) |  | .52 (.62) |
| cat8 | .87 (.70) | .68 (.65) | .52 (.62) |  |

(c) Cat, view $V_{90}$

|  | $gorilla_1$ | $gorilla_5$ | $gorilla_8$ | $gorilla_{14}$ |
|---|---|---|---|---|
| gorilla1 |  | .93 (.51) | .92 (.52) | .85 (.62) |
| gorilla5 | .93 (.51) |  | .94 (.46) | .74 (.55) |
| gorilla8 | .92 (.52) | .94 (.46) |  | .69 (.53) |
| gorilla14 | .85 (.62) | .74 (.55) | .69 (.53) |  |

(d) Gorilla, view $V_0$

|  | $gorilla_1$ | $gorilla_5$ | $gorilla_8$ | $gorilla_{14}$ |
|---|---|---|---|---|
| gorilla1 |  | .91 (.52) | .90 (.45) | .82 (.55) |
| gorilla5 | .91 (.52) |  | .88 (.51) | .85 (.47) |
| gorilla8 | .90 (.45) | .88 (.51) |  | .81 (.50) |
| gorilla14 | .82 (.55) | .85 (.47) | .81 (.50) |  |

(e) Gorilla, view $V_{45}$

|  | $gorilla_1$ | $gorilla_5$ | $gorilla_8$ | $gorilla_{14}$ |
|---|---|---|---|---|
| gorilla1 |  | .71 (.48) | .40 (.40) | .38 (.33) |
| gorilla5 | .71 (.48) |  | .72 (.38) | .45 (.34) |
| gorilla8 | .40 (.40) | .72 (.38) |  | .27 (.41) |
| gorilla14 | .38 (.33) | .45 (.34) | .27 (.41) |  |

(f) Gorilla, view $V_{90}$

|  | $horse_0$ | $horse_7$ | $horse_{10}$ | $horse_{17}$ |
|---|---|---|---|---|
| horse0 |  | .59 (.59) | .44 (.52) | .74 (.77) |
| horse7 | .59 (.59) |  | .39 (.57) | .34 (.56) |
| horse10 | .44 (.52) | .39 (.57) |  | .80 (.50) |
| horse17 | .74 (.77) | .34 (.56) | .80 (.50) |  |

(g) Horse, view $V_0$

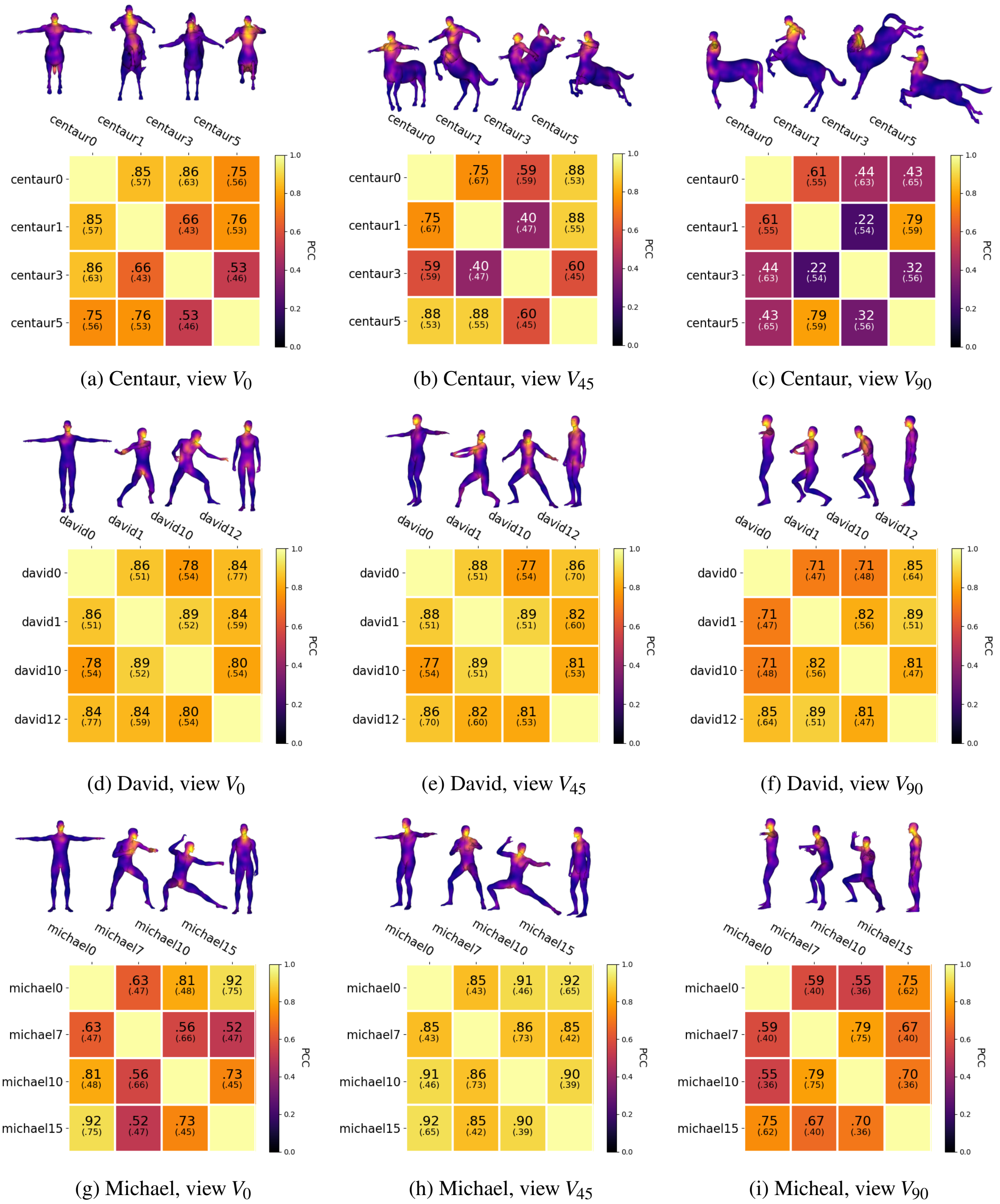
|  | $horse_0$ | $horse_7$ | $horse_{10}$ | $horse_{17}$ |
|---|---|---|---|---|
| horse0 |  | .73 (.80) | .83 (.79) | .82 (.80) |
| horse7 | .73 (.80) |  | .79 (.76) | .75 (.76) |
| horse10 | .83 (.79) | .79 (.76) |  | .82 (.74) |
| horse17 | .82 (.80) | .75 (.76) | .82 (.74) |  |

(h) Horse, view $V_{45}$

|  | $horse_0$ | $horse_7$ | $horse_{10}$ | $horse_{17}$ |
|---|---|---|---|---|
| horse0 |  | .77 (.76) | .87 (.84) | .84 (.77) |
| horse7 | .77 (.76) |  | .74 (.78) | .75 (.83) |
| horse10 | .87 (.84) | .74 (.78) |  | .84 (.80) |
| horse17 | .84 (.77) | .75 (.83) | .84 (.80) |  |

(i) Horse, view $V_{90}$

Figure 5: Similarity matrices for characters: *Cat*, *Gorilla*, and *Horse*. The similarity matrices show the Pearson's Linear Correlation (PCC) values between pairs of shapes belonging to the same character and looked from the same view. The corresponding Jaccard Index is indicated between parentheses. The aggregated 3D fixation maps related to each pose are shown above each similarity matrix. The 3D fixation maps are color coded for visualization purpose, with yellow representing the highest number of fixations and blue indicating areas with no fixations.



(a) Centaur, view $V_0$

(b) Centaur, view $V_{45}$

(c) Centaur, view $V_{90}$

(d) David, view $V_0$

(e) David, view $V_{45}$

(f) David, view $V_{90}$

(g) Michael, view $V_0$

(h) Michael, view $V_{45}$

(i) Micheal, view $V_{90}$

Figure 6: Similarity matrices for characters: *Centaur*, *David*, and *Michael*. The similarity matrices show the Pearson's Linear Correlation (PCC) values between pairs of shapes belonging to the same character and looked from the same view. The corresponding Jaccard Index is indicated between parentheses. The aggregated 3D fixation maps related to each pose are shown above each similarity matrix. The 3D fixation maps are color coded for visualization purpose, with yellow representing the highest number of fixations and blue indicating areas with no fixations.

Figure 3 shows the mean and standard deviation of the Jaccard Index values (intersection over union) computed from pairs of different poses of the same character. These values express a measure of variability of visible vertices between a pair of poses. Since PCC is computed on the intersection of the sets of visible vertices, the Jaccard Index measures also the extent of the portion of the mesh on which the corresponding PCC value was measured in relation to the union of the visible vertices. For example, the mean Jaccard Index for the gorilla shapes viewed from view $V_{90}$ is 0.39 which implies that on average the pairs of these 3D meshes share only about two-fifths of the union of the visible vertices from view $V_{90}$ due to variability in the poses. While the mean Jaccard Index value for the horse shapes viewed from view $V_{90}$ is 0.79 which indicates that on average the pairs of these 3D meshes share about four-fifths of the union of the visible vertices, denoting that very similar portions of the mesh are visible from that view.

All the computed PCC resulted in positive values, hence we treat the data as similarity values between 0 and 1. Figure 4 shows the mean and standard deviation values obtained by pairs of 3D fixation maps belonging to the same character and view.

For some of the characters, the mean PCC values show a low similarity indicating that a different pose might influence visual attention. A variation of mean values of the character between views reveals a notable impact also of the view itself. This trend is not constant between characters, revealing that some poses seen from specific views influence the visual attention more than others, as shown by the similarity matrices of each character and view in Figures 5 and 6.

Three examples of low similarity values can be found in the *Gorilla* and *Centaur* characters from view $V_{90}$ (Figures 5f and 6c) and the *Horse* character from view $V_0$ (Figure 5g). The corresponding 3D fixation maps show high fixation values over geometric salient features such as folds on the mesh (e.g. the fold between torso and leg in *Gorilla14*) as well as parts of the mesh related to the action the shape could represent, e.g. the front leg up in the air for *Horse7*.

Further examples indicating the influence of action representation can be found comparing the 3D fixation maps of specific pairs of poses. *Cat2* pose presents a cat licking its right paw, while *Cat8* shows a cat lurking. In Figures 5b and 5c, the 3D fixation maps of *Cat2* indicate considerable higher fixation values on the right leg compared to the 3D fixation maps of *Cat8*. This could be related to the proximity of the leg to the head of the cat or to the actual relevance of the leg in the context of the action that is represented. In Figure 6b, *Centaur1* and *Centaur3* show two opposite poses, the first with the front legs raised off the ground, while the second with the back legs up in the air. In both cases, the 3D fixation maps indicate higher fixation values corresponding to the pair of legs raised off the ground, the agent of the represented actions. It appears that the fixation patterns for some characters are clustered on areas related to a potential action which leads to additional questions to explore further.

## 9 CONCLUSIONS

In this work, we presented a first investigation of the impact of different poses of 3D meshes on gaze data. The aggregated fixation data of the participants obtained by an eye-tracker were projected onto the 3D meshes. Pearson's Correlation Coefficient was used as a measure of similarity between pairs of 3D fixation maps of the same character in different poses and from different views. The obtained 3D fixation maps are coherent with previous studies in that fixation data are focused strongly on the facial regions [How05, McD09, Lav18]. In addition, this paper further indicates that the fixations on the facial regions also appear on characters in different poses. The PCC results show low similarity values between different poses for some of the tested characters. In particular, the low similarity between some of the poses appeared linked with fixations focusing on two factors: geometric salient features and semantically salient parts caused by potential actions or gestures. These results indicate the necessity of further investigations. For example, it would be relevant in future research to explore further the relationship between these two factors and to investigate the influence of different types of actions.

## 10 ACKNOWLEDGMENTS

This work was supported in part by KK-stiftelsen Sweden, through the ViaTecH Synergy Project (contract 20170056).

## 11 REFERENCES

[Bor13] Borji, A., and Itti, L. State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence 35*, 1 (Jan. 2013), 185–207.

[Bro08] Bronstein, A., Bronstein, M., and Kimmel, R. *Numerical Geometry of Non-Rigid Shapes*, 1 ed. Springer Publishing Company, Incorporated, 2008.

[Byl19] Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., and Durand, F. What Do Different Evaluation Metrics Tell Us About Saliency Models? *IEEE Transactions on Pattern Analysis and Machine Intelligence 41*, 3 (March 2019), 740–757.

[Cun11] Cunningham, D. W., and Wallraven, C. *Experimental Design : From User Studies to Psychophysics*. CRC Press LLC, 2011.

[Duc17] Duchowski, A. T. *Eye Tracking Methodology: Theory and Practice*, 3rd ed. Springer Publishing Company, Incorporated, 2017.

[How05] Howlett, S., Hamill, J., and O'Sullivan, C. Predicting and Evaluating Saliency for Simplified Polygonal Models. *ACM Transactions on Applied Perception 2*, 3 (July 2005), 286–308.

[Kim10] Kim, Y., Varshney, A., Jacobs, D. W., and Guimbretière, F. Mesh Saliency and Human Eye Fixations. *ACM Transactions on Applied Perception 7*, 2 (Feb. 2010), 12:1–12:13.

[Kom10] Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., and Gowda, S. M. Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *IEEE Transactions on Biomedical Engineering 57*, 11 (Nov. 2010), 2635–2645.

[Lar11] Larkin, M., and O'Sullivan, C. Perception of Simplification Artifacts for Animated Characters. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (New York, NY, USA, 2011), APGV'11, ACM, pp. 93–100.

[Lav18] Lavoué, G., Cordier, F., Seo, H., and Larabi, M.-C. Visual Attention for Rendered 3D Shapes. *Computer Graphics Forum 37*, 2 (2018), 191–203.

[Lee05] Lee, C. H., Varshney, A., and Jacobs, D. W. Mesh Saliency. In *ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), SIGGRAPH '05, ACM, pp. 659–666.

[Lei16] Leifman, G., Shtrom, E., and Tal, A. Surface Regions of Interest for Viewpoint Selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence 38*, 12 (Dec. 2016), 2544–2556.

[Mes] Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., and Ranzuglia, G. MeshLab: an Open-Source Mesh Processing Tool. In *Eurographics Italian Chapter Conference* (2008), V. Scarano, R. D. Chiara, and U. Erra, Eds., The Eurographics Association.

[McD09] McDonnell, R., Larkin, M., Hernandez, B., Rudomin, I., and O'Sullivan, C. Eye-catching Crowds: Saliency Based Selective Variation. *ACM Transactions on Graphics 28*, 3 (July 2009), 55:1–55:10.

[Par02] Parkhurst, D., Law, K., and Niebur, E. Modeling the Role of Salience in the Allocation of Overt Visual Attention. *Vision Research 42*, 1 (2002), 107–123.

[Pat16] Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D., and Lefohn, A. Towards Foveated Rendering for Gaze-Tracked Virtual Reality. *ACM Transactions on Graphics 35*, 6 (Nov. 2016), 179:1–179:12.

[Tor06] Torralba, A., Oliva, A., Castelhano, M. S., and Henderson, J. M. Contextual Guidance of Eye Movements and Attention in Real-world Scenes: The Role of Global Features in Object Search. *Psychological Review 113*, 4 (2006), 766–786.

[Tos] TOSCA High-Resolution Dataset. `http://tosca.cs.technion.ac.il/book/resources_data.html`.

[Wan16] Wang, X., Lindlbauer, D., Lessig, C., Maertens, M., and Alexa, M. Measuring the Visual Salience of 3D Printed Objects. *IEEE Computer Graphics and Applications 36*, 4 (July 2016), 46–55.

[Wan18] Wang, X., Koch, S., Holmqvist, K., and Alexa, M. Tracking the Gaze on Objects in 3D: How Do People Really Look at the Bunny? *ACM Transactions on Graphics 37*, 6 (Dec. 2018), 188:1–188:18.

[Wu13] Wu, J., Shen, X., Zhu, W., and Liu, L. Mesh Saliency with Global Rarity. *Graphical Models 75*, 5 (2013), 255–264.