

Perceptual Adjustment of Eyeball Rotation and Pupil Size Jitter for Virtual Characters

SOPHIE JÖRG, Clemson University

ANDREW T. DUCHOWSKI, Clemson University

KRZYSZTOF KREJTZ, SWPS University of Social Sciences and Humanities

ANNA NIEDZIELSKA, SWPS University of Social Sciences and Humanities

Eye motions constitute an important part of our daily face-to-face interactions. Even subtle details in the eyes' motions give us clues about a person's thoughts and emotions. Believable and natural animation of the eyes is therefore crucial when creating appealing virtual characters. In this paper, we investigate the perceived naturalness of detailed eye motions, more specifically of jitter of the eyeball rotation and pupil diameter on three virtual characters with differing levels of realism. Participants watched stimuli with six scaling factors from 0 to 1 in increments of 0.2 varying eye rotation and pupil size jitter individually and had to indicate if they would like to increase or decrease the level of jitter to make the animation look more natural. Based on participants' responses, we determine the scaling factors for noise attenuation perceived as most natural for each character when using motion-captured eye motions. We compute the corresponding average jitter amplitudes for the eyeball rotation and pupil size to serve as guidelines for other characters. We find that the amplitudes perceived as most natural depend on the character with our character with a medium level of realism requiring the largest scaling factors.

CCS Concepts: • **Computing methodologies** → **Animation; Perception.**

Additional Key Words and Phrases: character animation, eye motion, perceptual study, jitter, pupil dilation

ACM Reference Format:

Sophie Jörg, Andrew T. Duchowski, Krzysztof Krejtz, and Anna Niedzielska. 2018. Perceptual Adjustment of Eyeball Rotation and Pupil Size Jitter for Virtual Characters. *ACM Trans. Appl. Percept.* 1, 1, Article 1 (January 2018), 14 pages. <https://doi.org/0000001.0000001>

1 INTRODUCTION

In our daily face-to-face interaction, we do not only communicate through spoken words, but our face and body motions play important roles in conveying information. As a subset of facial movements, eye motions have always been seen as a critical constituent of this communication. The eyes have been called a window to the soul and the interpreter of the mind [36]. Other people's gaze gives us indications about their thoughts, their emotions, and their mood. Missing empathy for an animated character or uncanny feelings have been attributed to cold and "dead-eyed" characters [13, 21, 24]. There is no doubt that eye animation is crucial when creating an appealing virtual character.

From an animation point of view, eye motions consist of several components: the eyeball motions, which entail fixations and saccades; the changes in size of the pupil diameter; and periocular motions mainly consisting of

Authors' addresses: Sophie Jörg, School of Computing, Clemson University, sjoerg@clemson.edu; Andrew T. Duchowski, School of Computing, Clemson University, duchowski@siggraph.org; Krzysztof Krejtz, Psychology Department, SWPS University of Social Sciences and Humanities, kkrejtz@swps.edu.pl; Anna Niedzielska, Psychology Department, SWPS University of Social Sciences and Humanities, aniedzielska@st.swps.edu.pl.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

1544-3558/2018/1-ART1

<https://doi.org/0000001.0000001>

blinks and lid saccades. The creation and perception of larger and more obvious eye motions such as saccades and smooth pursuits have been widely investigated [19, 36, 44]. Eye blinks have also been well characterized [39, 41].

The focus of this paper is on more subtle eyeball motions and pupil diameter changes. Duchowski et al. [6, 7] and Krejtz et al. [17] have shown that the inclusion of procedural eyeball rotation and pupil size jitter is preferred to no jitter at all using a simple black and white character model. However, motions recorded with professional eye-tracking equipment contain measurement noise that is perceived to be overly jittery [7]. While this noise is typically filtered during eye motion analysis, it is visible when directly applied to a virtual character. Smoothing is a standard procedure when post-processing motion captured movements. Typical smoothing methods such as Gaussian filters would not only decrease the noise in our recordings but also reduce the quality and accuracy of the eye motion signal.

In previous work, eye-tracking data has been scaled down, which not only affects the noise but also changes the larger saccadic motions [6]. In this work we suggest to first detect saccades and fixations in the unprocessed eye movement data and then to attenuate only the noise within the fixations. Furthermore, the amplitudes of eyeball rotation and pupil size jitter that are perceived as most natural on a virtual character have never been thoroughly investigated: the ideal scaling factors are unknown.

In this paper we approach the problem empirically and determine preferred amplitudes of eyeball rotation and pupil size jitter with a perceptual experiment. The values perceived as most natural might depend on the virtual character. These amplitudes might furthermore not correspond to the motions that people exhibit naturally, especially for characters that are less realistic. We therefore investigate the amplitudes of eyeball rotation and pupil size jitter on the perceived naturalness of the eye movements of characters at different levels of realism. The main questions we address are:

- Which amplitudes of eyeball and pupil size jitter are perceived as most natural on a virtual character?
- Does perceived naturalness of simulated natural jitter depend on the realism of the character model?

To answer these questions, we record and post-process eye motions to create our stimuli and conduct our study on three virtual characters: a very realistic virtual human, a less realistic human-like character, and an anthropomorphic robot (see Figure ??).

Our results suggest that the scaling parameters required for fine-tuning motion-captured data via signal attenuation depend on the animated character. As noise levels of different eye-trackers of course differ [43], we compute a meaningful representation of these parameters for future reference for animators.

In the following sections, we summarize previous work in this area and clarify our contributions (Section 2). We then give a detailed account of our method (Section 3), including the stimuli creation, and of our results (Section 4). We discuss our outcomes in Section 5 and conclude with a summary and ideas for future work in Section 6.

2 RELATED WORK

In this work, we take a closer look at the very subtle motions of the eyeball and pupil. It has been shown that viewers are able to perceive and interpret very subtle human motions [16], anomalies in facial motions being particularly salient [13]. Noise has been found to increase the perceived naturalness of eye animations [6, 17]. Further work evaluates how changes in eye motions affect perception, e.g., how trust is conveyed through gaze [30]. Concerning eye motions specifically, even just changes in the periocular motions can affect how the personality of a character is judged: the blink rate of a virtual character has an effect on the friendliness and intelligence attributed to it [40].

However, it is very difficult to record the subtle motions of the eye accurately as “tremor amplitudes and frequencies are usually in the range of the recording system’s noise.” [25]. Most approaches to capture eye motions are either image-based or use IR illumination. While progress in image-based methods now allows for real-time

gaze animation even with variations in lighting [42], the details of fixational movements are not captured with these techniques.

A large body of research has investigated how to create high-level gaze behavior such as saccades or smooth pursuits for virtual characters [2, 5, 19, 23, 31, 44]. Ruhland et al. [36] survey these works in a state-of-the-art report detailing research on saccades, the vestibulo-ocular reflex, smooth pursuit, vergence, eyelid movement, and combined eye-head movements followed by work on higher-level aspects of gaze behavior. The more subtle pupil size and eyeball rotation jitter are not taken into account. More current work has shown that adding some noise to the eyeball rotation and pupil size increases the perceived naturalness of virtual characters [6, 17]. However, in this work a level of noise perceived as natural was guessed and eye tracking data was roughly halved accordingly. It has not been investigated systematically, which average amplitudes of noise provide motions perceived as most natural, which is the contribution of our work.

For comparison to our results, we describe real eye movements. Saccades are fast jumps of both eyes: the eyeballs are accelerated very quickly to look at a new gaze point within tens of milliseconds. During fixations, when we look at a fixed point between saccades, our gaze is nearly but not quite stationary. Specifically, three types of subtle eye motions occur: microsaccades, ocular drift, and ocular microtremor [33]. Only recently have these motions and their functions been analyzed in more detail [26].

Microsaccades are jerk-like motions similar to saccades. Their amplitude has been measured between 2 and 50 arcminutes with frequencies around 1 Hz. Drifts are slower motions that might reflect a random-walk-like process [1]. Finally, ocular microtremors have frequencies around 90 Hz, can last up to 1 second and have been measured with amplitudes between 1 and over 30 arcminutes. The role of these subtle motions is still being debated. Some of them might result from noise in the neuronal firing to the muscles. There is also evidence that they have an important role in our perception: if they are counteracted, our visual perception fades away [25].

Pupil size is well known to react to illumination [20]. The pupil dilates in the dark and constricts in bright light. Pupil diameter typically varies between 3 and 7.5 mm in healthy adults. However, similar to the eyeballs, the pupils are never completely still. They show slow waves of changes in size with amplitudes up to ± 0.5 mm and a duration between 4 and 40 seconds and faster, smaller oscillations with variations of 0.1 to 0.3 mm that take 0.5 to 1 second [22]. Pupil size oscillations, also called pupil unrest [38], are furthermore influenced by a range of mental processes and physical characteristics such as attention, arousal, fatigue, and age [4, 15, 18, 22, 29].

While these motions are very subtle, animators are aware that even surprisingly detailed movements might be perceived and may convey an emotion, mood, or intention, and can be crucial when bringing to life a virtual character [37]. The possible level of realism of virtual characters might contribute to our ability to detect and interpret subtle details. Recent work shows how to create virtual character models that can barely be distinguished from reality. Bérard et al. [3] recently presented a parametric model to reconstruct eyeball geometries and textures using a database of high-frequency scans. The results look highly realistic. The subtle motions of eyes have also been used to simulate our perception of our environment. Ritchel et al. [35] propose a model to simulate dynamic glare that is created by light scattered within the human eye when looking at bright light sources. With increasing realism and attention to detail of such models it becomes even more important to systematically investigate our perception of more subtle details in eye animations.

Results on the effect of character realism on the perception of gaze are mixed. Ruhland et al. [36] find that the perception of personality remains robust across two characters of different realism. Garau et al. [11, 12] found a strong interaction effect between the realism of the character and the gaze type: for the more realistic character, more elaborate gaze behavior was preferred, whereas for the less realistic character, random gaze patterns received better ratings. This finding corresponds to studies using the whole body that have shown that the realism of a virtual character or even just its render style might affect how we perceive motions [14, 27, 28].

Based on these previous findings, our contributions are as follows: We provide details on how eye-tracking data can be processed to make it suitable for eye movement animation by attenuating the noise in the signal

without affecting the saccadic motions. We present the first study that systematically finds the scaling values perceived as most natural for this purpose for three virtual characters with different levels of realism. While we found in a perceptual experiment that the detailed scaling values we used in our approach need to be adjusted for each character, we provide average amplitudes for eyeball rotation jitter and pupil size jitter that can be used as reference values.

3 METHOD

The goal of our experiment is to find jitter amplitudes for the eyeball rotation and for the pupil diameter that are perceived to be most natural on virtual characters. We hypothesize that the resulting values differ for virtual characters with different levels of realism.

3.1 Experiment Design

Our experiment uses a $2 \times 6 \times 3 \times 2$ mixed design with the within-subject variables Component, Scaling factor, and Character and the between-group variable Motion. We roughly followed a staircase procedure, adjusting it based on our pilot studies.

For the purpose of this study we recorded two distinct eye motions from a person and attenuated the noise of the data. We then displayed these motions on three characters with different levels of realism. In summary, our stimuli varies based on four factors:

Component : we influenced the amount of jitter in two eye motion components, eyeball rotation or eye jitter (EJ) and pupil size or pupil dilation (PS)

Scaling factor : we varied the jitter amplitude by attenuating our captured motions with six different scaling factors, see below

Character : we created three different character models, a highly realistic human (H), a less realistic and more cartoony human (C), and an anthropomorphic robot (R)

Motion : two motions were recorded: a nine point calibration-like sequence and reading (see above)

The scaling factors used for eye jitter (EJ) were 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0. The scaling factor is the value with which the difference between a data value and the mean of the values within a fixation is being multiplied. So a scaling factor of 0.0 means that the difference to the average value within the fixation has been eliminated which results in constant values (no jitter at all) and a scaling factor of 1.0 represents the original motion captured data. Further details are described at the end of Section 3.2.2. The scaling factor step size of 0.2 was chosen as the differences it creates are subtle but still visible in direct comparison. For the pupil size (PS), the differences to the average values were scaled down in a similar way, but for the whole recording at once. The same 6 values from 0.0 to 1.0 were used as scaling factors. Based on preliminary tests, we expected both the pupil size scale factor and the eye jitter scaling factor to be around 0.5 or slightly lower. A value of 0.5 was therefore assigned to the non-varying scaling factor for all stimuli in this experiment. In total, the experiment comprises 2 (EJ and PS) $\times 6$ (scaling factors) $\times 3$ (characters) $\times 2$ (motions) = 72 animations. In the following section we describe this process in more detail.

3.2 Stimuli Creation

3.2.1 Capturing Eye Motions. Our aim was to record two motions with different characteristics. We chose two tasks: looking at nine points on a plane perpendicular to the viewer, and reading (also on a plane perpendicular to the viewer; see Figure 1).

The nine point task requires the participant to follow a sequence of numbers shown in Figure 1(a) similar to what is done during a standard eye-tracker calibration procedure. The participant fixated on each number for two

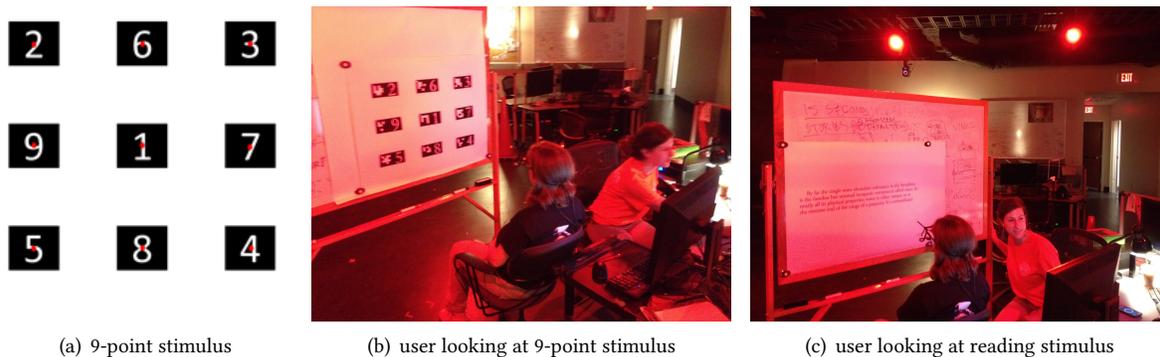


Fig. 1. Motion capture setup to record eye movement data when looking at a 9-point grid and when reading a short passage. We later added a headrest to our setup to increase the accuracy of our measurements.

seconds. A timed chime informed the participant to proceed to the next number. The task was repeated to allow for familiarization with the sequence. This task produces a mix of relatively long fixations and large saccades.

The reading task asks the participant to read a four line excerpt of text used in an eye-tracking study by [34]:

*By far the single most abundant substance in the biosphere
is the familiar but unusual inorganic compound called water. In
nearly all its physical properties water is either unique or at
the extreme end of the range of a property. It's extraordinary.*

Reading leads to shorter fixations and smaller saccades and is a pattern that has been analyzed and described in much detail. The nine point animation is 20 seconds long, the reading motion 12seconds.

As eye motion components can vary with emotion, we chose neutral stimuli. The gaze point distance was kept constant to avoid any changes in vergence: the points and the text were placed on a whiteboard at a fixed distance to the participant (approximately 78.5 cm). The eye motions for the two experiments were recorded with a Dikablis eye tracker from Ergoneers, a head-mounted unit operating at 25 Hz, which, according to the manufacturer, measures gaze direction with an accuracy of 0.5 degrees visual angle. The head was tracked with a Vicon motion capture system. The systems were integrated using the software D-Lab by Ergoneers and Tracker by Vicon. For added accuracy a headrest was used during recording. The collected data includes gaze positions and pupil diameter.

3.2.2 Processing the Eye-tracked Data Signal. To attenuate eye tracker noise, we detect fixations in the unprocessed eye movement data provided by the eye tracker (by unprocessed we mean unfiltered by the eye tracker's software). We then attenuate the data within each fixation. With this technique, we reduce the noise within the fixations without affecting the larger signals.

To detect fixations we first detect saccades, based on a variant of Engbert and Kliegl's 2003 velocity-based algorithm for detection of microsaccades in two-dimensional velocity space. The algorithm differentiates the gaze position signal $\mathbf{p}_t = (x(t), y(t))$ to obtain velocity which is then thresholded to delineate between saccades (when the eyes move very quickly) and the remaining signal which is presumed to be captured when gaze is fixed on a stationary object (during a fixation).

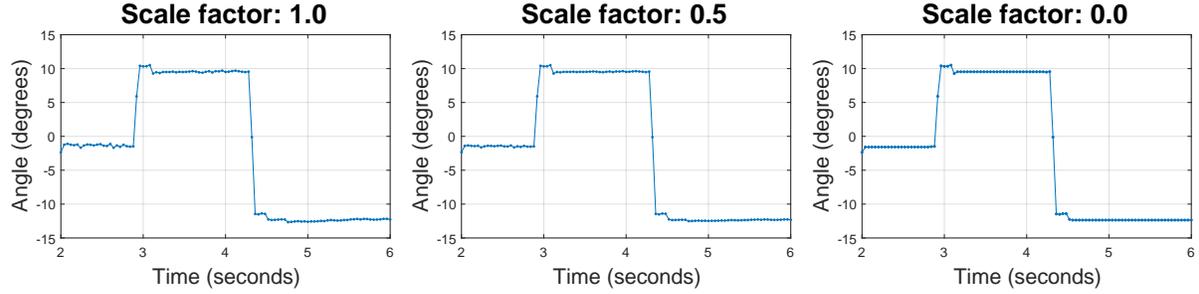


Fig. 2. Effect of scaling factors on gaze x -coordinate over time. The gaze point values are only attenuated during the fixations thus not erasing the detailed dynamics during saccades.

The algorithm proceeds in three steps. First, we transform the time series of gaze positions to velocities via a moving average of velocities over 5 data samples,

$$\dot{x}_n = \frac{x_{n+2} + x_{n+1} - x_{n-1} - x_{n-2}}{6\Delta t}, \quad (1)$$

computed separately for $x(t)$ and $y(t)$. As Engbert and Kliegl note, due to the random orientations of the velocity vectors during fixation, the resulting mean value is close to zero. Saccades, being ballistic movements, create small linear sequences that can be identified by their velocities.

Second, computation of velocity thresholds for the detection algorithm is based on the median of the velocity. A multiple of the standard deviation of the velocity distribution is used as the detection threshold [8],

$$\sigma_x = \sqrt{\langle \dot{x}^2 \rangle - \langle \dot{x} \rangle^2}, \quad \sigma_y = \sqrt{\langle \dot{y}^2 \rangle - \langle \dot{y} \rangle^2}$$

where $\langle \cdot \rangle$ denotes the median estimator. Detection thresholds are computed independently for horizontal η_x and vertical η_y components and separately for each trial, relative to the noise level, i.e., $\eta_x = \lambda\sigma_x$, $\eta_y = \lambda\sigma_y$.

Third, we tailor our version of Engbert and Kliegl's [9] algorithm by adjusting λ to detect large amplitude saccades as well as saccades with smaller amplitudes. The idea is to preserve the saccades and only scale the data within fixations. Thus, unlike [6] who attenuate the whole signal, we only attenuate the signal within fixations, thereby retaining the distances between fixation points indicated by the saccades. Scaling the whole signal necessarily also scales the saccades, which reduces the apparent rotation of the eye, reducing the character's effective field of view. Large amplitude saccades are detected by setting $\lambda=23$; saccades with small amplitudes are detected by setting $\lambda=3$. This two-step detection process preserves saccades of amplitudes as small as about one degree (see Figure 2).

Once the saccades are detected, we compute the average \bar{x} of all points within each fixation. For each fixation, we then scale the differences between the gaze point $x(t)$ and the fixation mean \bar{x} by a scaling factor $s \in [0, 1]$ so that the scaled gaze position $x'(t)$ is:

$$x'(t) = \bar{x} + s(x(t) - \bar{x}) \quad (2)$$

Using $s \in [0, 1]$ gives us good control over simulated jitter within each fixation without disturbing the amplitude of saccades. The objective of our experiment is to empirically find suitable scaling values for s via perceptual evaluation of the produced eye motions (see Figure 2).

3.2.3 Character Models. Previous work has shown that motions can be perceived differently depending on the realism of the used virtual character [11, 12]. We therefore used three characters at different levels of realism: a highly realistic human character, a less realistic human character, and an anthropomorphic robot (see Figure ??).

The highly realistic human model was purchased and then rigged by students in digital arts to control the different aspects of eye motion. The other two character models were created and rigged by students. All eye colors were adjusted to allow for good contrast with the pupil.

Gaze, or look points, control where the eyeballs are looking in a 2-dimensional plane. The distance of the look points from the character was adjusted for each character to create results that look visually as similar as possible based on reference close-up videos of one eye that were recorded at the same time as the data used in this study was collected. Pupil dilation can be scaled based on its relative size compared to the iris. Both the upper and lower eyelids were rigged to follow the vertical motion of the eyeball as well as to blink from this adjusted position. The robot character's eyes were loosely inspired by the animated movie *Robots* released in 2005 and directed by Chris Wedge. They were designed to be cartoony in nature but to structurally remain similar to the setup for the human models. So the eyelids were designed as mechanical shutters that extend vertically to cause the character to blink.

3.2.4 Animation Creation. Gaze positions and pupil diameters were mapped onto our virtual characters. The motion of one eye was used to steer both eyes. Lid motions were not measured directly, instead the lids moved in response to modeled blinks and eyeball rotations. Lid closure during blinks was modeled by the function introduced by Duchowski et al. [6], who based it on the work of Trutoiu et al. [41]. The function results in an asymmetric blink with a fast down-phase when the upper lid lowers and a slower up-phase when the lid rises, corresponding to observed eyelid velocity profiles [10, 41].

Duchowski et al.'s blink function allows for specification of arbitrary blink duration [6]. When using motion-captured data, a blink is inferred from gaps in data which occur due to the eye tracker losing the image of the eye. Consequently, a sequence of frames is found in the data with null information for gaze position and pupil diameter. Blink animation is adjusted to the length of this gap.

Saccadic lid movements closely follow the saccadic movements of the eyeballs. We animate lid saccades by rigging the eyelids to be controlled by the vertical eyeball motion. Saccadic lid movements are animated at all times except during a blink motion. Throughout stimuli creation, animations were compared to ground truth close-up video capture of the recorded eye and across characters to ensure accurate correspondences.

We rendered extreme close-up views of each character and motion using SolidAngle's Arnold renderer. The extreme close-up was chosen to ensure that the eye motion details could be clearly seen and that participants would not be distracted by any facial animation or lack thereof.

3.3 Procedure

To investigate all factors we needed at least two staircases for each combination of character and motion for eye jitter and pupil size. The motions and the differences between them are very subtle and require concentration from participants. We therefore kept the participation time short and split the study in two experiments: one investigating the scaling factors in the nine point motion, the other examining reading motion. Each experiment used two blocks, one varying the pupil size scaling factor and one varying the eye jitter scaling factor. In both cases, animations were presented in random order with some participants first viewing the variations in eye jitter and others first viewing changes in the pupil size scaling factor. The variation of eye jitter and pupil size were not intermingled so that the participants could focus their attention on one factor at a time.

Participants were informed about the procedure of the experiment and filled out a consent form. It was explained that eye jitter or pupil size would be varied and they were asked to press the up arrow if they felt that more jitter would look more natural or the down arrow if they felt that less jitter would look more natural. The next animation then showed a higher or lower jitter following the participant's answer. If that was not possible because the bounds of 0.0 and 1.0 were reached, the same animation was played again.

The animations were looped and could be watched as long as the participants wanted but had to be watched for a minimum time of 10 seconds each. After 10 seconds, a prompt appeared allowing the up or down arrow keys to be pressed. In our pilot experiments, participants got frustrated that there was no option to express that they felt that the current jitter was natural, which led us to include the option of pressing the space bar after 5 trials to indicate that the current jitter was most natural. The following situations terminated the staircase and started the next: pressing the space bar, reversing the staircase more than four times in a row, wishing to increase the jitter at a scaling factor of 1.0 three times in a row, wishing to reduce the jitter at a scaling factor of 0.0 three times in a row, or watching 15 animations.

Two staircases were shown for each character, one starting with the highest scaling factor and one starting with the lowest scaling factor. This led to a total of 3 (character models) \times 2 (starting condition) \times 2 (eye jitter, EJ and pupil size, PS) = 12 staircases in each part of the study. A progress bar at the top of the screen indicated the progress throughout the experiment. We debriefed participants and asked for comments at the end of the experiments.

3.4 Participants and Apparatus

Eighteen participants (12 female and 6 male) between 19 and 45 years of age ($M = 24.78$, $SD = 7.46$) from the SWPS University of Social Sciences and Humanities in Warsaw, Poland, volunteered to take part in the first part of this study investigating the nine point motion in exchange for academic credit points. Participants were seated in front of a 20 inch screen. The screen resolution was 1440×900 .

The reading animations were viewed by 20 volunteers (11 female, 9 male) at the Interactive Technologies Laboratory in the National Information Processing Institute in Warsaw, Poland. Participants were between 26 and 43 years of age ($M = 35.35$, $SD = 4.93$). Participants were seated at about 57 cm of a 24" screen with 1920×1200 resolution (see Figure 3(a)). Both experiments were conducted using PsychoPy [32]. All participants were naïve as to the purpose of the experiment.

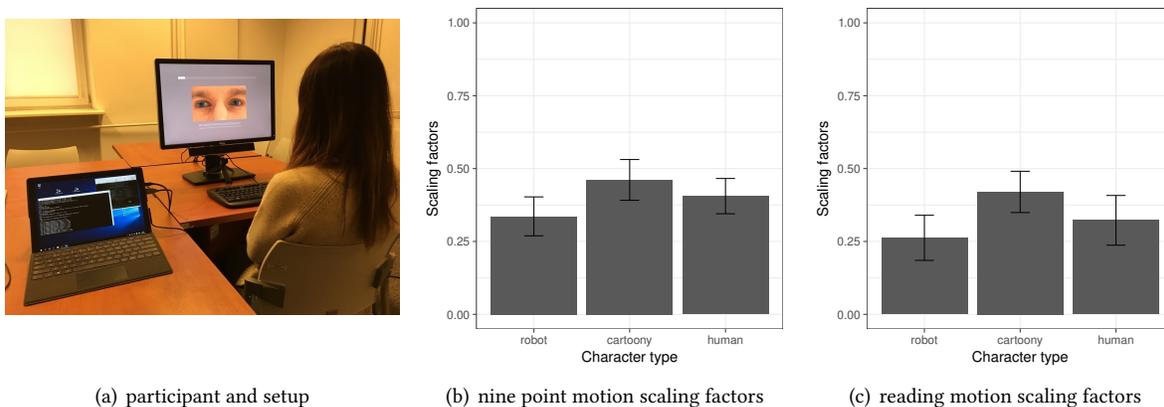


Fig. 3. Participant and experimental setup of our study, with averaged scaling factors perceived as most natural for the reading and nine point motion animations. Whiskers represent 95% confidence intervals.

4 RESULTS

The resulting scaling factor of each staircase and each participant depends on the termination condition. When the space bar was pressed to indicate that the current level of jitter seemed most natural, the scaling factor

Table 1. Averaged scaling factors perceived as most natural for the eye motion jitter / Corresponding average jitter amplitudes of eye eyeball rotation expressed in degrees.

motion	highly realistic character (H)	less realistic character (C)	anthropomorphic robot (R)
nine point motion	0.43 / 0.14	0.48 / 0.10	0.32 / 0.11
reading motion	0.28 / 0.12	0.37 / 0.07	0.28 / 0.09

Table 2. Averaged scaling factors perceived as most natural for the pupil size jitter / Corresponding average amplitudes of the jitters affecting the pupil size in proportion of the iris diameter.

motion	highly realistic character (H)	less realistic character (C)	anthropomorphic robot (R)
nine point motion	0.38 / 0.0023	0.44 / 0.0026	0.35 / 0.0021
reading motion	0.37 / 0.0013	0.48 / 0.0016	0.25 / 0.0011

corresponding to that jitter was used as the result. When the staircase was terminated because the bounds were reached, the appropriate highest possible value (1.0) or lowest possible value (0.0) was used. Finally, when the staircase was reversed more than four times in a row, the last four scaling factors were averaged.

The last termination condition, watching 15 animations, did not occur in our study. From both experiments, the average number of videos watched per staircase was 5.2, the minimum being 3 and the maximum 12. The most frequent reason for ending a staircase was pressing the space bar (73.5%). The next most frequent reason was reaching the lower bound (19.5%) and then the higher bound (6.4%). Four reversals only happened three times.

Tables 1 and 2 list the averaged scaling factors perceived as most natural the eye motion and pupil size jitters, respectively. We furthermore list the average amplitudes of the jitters in angular values for the eyeball rotations and in proportion of iris size for the pupil size. The translation from scaling factors to average jitter amplitude is not as trivial as it might first seem. Each character has its own geometry with its own eye shape. For example, the robot's eyes are similar to ellipsoids. They are not rotating, but the gaze points are directly controlling the motion of the iris and pupil. The adjustments between characters to create visually similar results cause different translations from scaling factors to angular values for each character. Furthermore, as the nine point motion and the reading motion did not have the same amount of jitter in the originally motion captured motion, the correspondence of scaling factors to angular values also differs between the two scenes. We computed the average angular jitter amplitude based on the look points and the distance of the lookpoints to the eyes, using eyeball centers as a reference. We furthermore report the average amplitudes of the jitters affecting the pupil size. They are represented in proportion of the iris diameter, 1 meaning that the pupil is the same size as the iris (around 12 mm for an average adult) and 0 being a point. To give the reader a comparison: the pupil size in all our animations varies between 0.4 and 0.5. The values demonstrate how subtle our changes are.

Below, we analyze the differences for statistical significance between scaling factors, if any.

4.1 Nine point animation scaling factors

For the nine point motion experiment data, we performed a two-way (2×3) within-subjects ANOVA of scaling factors chosen by participants as the most natural, with jitter (eye motion vs. pupil size) and character animation (highly realistic (H) vs. less realistic (C) vs. anthropomorphic robot (R)) as the fixed factors. The analyses showed a significant main effect of character type, $F(1.55, 26.27) = 10.65$, $p < 0.001$, $\eta^2 = 0.05$, see Figure 3(b). The following

pairwise comparisons revealed, as for reading motion data analyses, that, to be perceived as natural, the cartoony character animation required a significantly greater scaling factor ($M=0.46$, $SE=0.04$) than the robot character ($M=0.34$, $SE=0.03$) or human character animations ($M=0.41$, $SE=0.03$). Note, that neither main effect of jitter $F(1, 17) < 1$, nor interaction between character and jitter, $F(1.47, 25.03) < 1$, was statistically significant.

4.2 Reading animation scaling factors

A two-way ANOVA, with (2×3) design analogous to the above, was also performed on the scaling factor perceived as most natural for reading motion animations. The analysis revealed a statistically significant main effect of animated character, $F(1.85, 35.23) = 7.95$, $p < 0.001$, $\eta^2 = 0.05$, see Figure 3(c). The following pairwise comparisons showed that, to be perceived as natural, the cartoony character animations required a significantly greater scaling factor ($M=0.42$, $SE=0.04$) than the robot character ($M=0.26$, $SE=0.04$) or the human character animations ($M=0.32$, $SE=0.04$). The difference between scaling factors of robot and human character animations was not statistically significant. Note that neither main effect of jitter, $F(1, 19) < 1$, nor interaction between character and jitter, $F(1.87, 35.60) = 1.90$, $p = 0.17$, was statistically significant.

5 DISCUSSION

We find that fine-tuning motion-captured data via signal attenuation is dependent on the animated character. Each scaling needs to be adjusted based on the rig and the appearance of the character. The angular values of the average eyeball jitter and the average pupil size jitters that we found can be used by animators as a reference. Interestingly, these values from 0.07 to 0.14 degrees (or 4.2 to 8.4 arcminutes) are within the measurements that have been done concerning the amplitudes of microsaccades and microtremors (between 2 and 50 arcminutes and 1 and 30 arcminutes, respectively). The jitter amplitudes perceived as most realistic for pupil size oscillations are smaller than what we found in the literature. However, the described oscillations are also much slower and do not correspond to what could be seen as noise in the neuronal firing to the muscles.

While it is interesting to observe that our results for the eyeball rotation jitter fall within the range that corresponds to real motions, a different result could also have been possible as our experiment measured the motion *perceived* as most natural and the stimuli does not present situations close to those typically happening in our daily lives. We rarely stare at other people's eyes from a very close distance for several seconds. We still believe that the averages we determined provide valuable information to animators, even if they should of course use their best judgment to create convincing animations for each character in each situation.

In the staircase procedure, in 19.5% of the cases reaching the lower bound was the reason to terminate the staircase. While we can not reduce the jitter to less than none, this indicates that some participants might prefer a complete lack of jitter in some cases. A closer look at individual responses from participants gives us further insights: If considering participants who chose a lack of jitter at least 5 times out of 6 responses, reaching either the lower bound or pressing the space bar as a termination condition, in the nine point motion experiment three participants preferred a complete lack of pupil size jitter whereas no participant consistently opted for no eyeball jitter. When watching the reading motion 6 participants preferred no jitter at all for the eyeballs, 2 viewers favored a lack of pupil size jitter with one participant present in both categories. But most participants did prefer some level of jitter. A typical example would be P11 who watched the nine point motions and chose the following scaling values: 0.2 twice, 0.4 twice, and 0 and 0.4, for the realistic human, the cartoony human, and the robot, respectively, when changing the pupil size jitter and 0.4 and 0.6 for each of the three characters when affecting the eye gaze jitter. Only one participant would have consistently preferred a larger jitter than what our animations provided. That participant was viewing the reading motions and gave these answers only for the eyeball jitter.

The cartoony character required the largest scaling factors and different assumptions could explain and interpret this result. Garau et al. indicated that viewers prefer more elaborate gaze behavior with more realistic characters [11, 12], which might explain why less jitter was required for the robot. As this work is several years old and the quality of the characters at the time was lower, we can not draw conclusions related to the highly realistic virtual human from it. It could also be the case that a realistic robot, the way most people imagine it, might not have any eyeball rotation or pupil size jitter. During the debriefing of our experiment, two participants expressed that the difference between the animations in the staircase procedure was more conspicuous for the highly realistic virtual human than for other characters. This observation might explain why the scaling factors for the most realistic character were lower than for the cartoony character.

6 CONCLUSION AND FUTURE WORK

In this paper, we create and perceptually evaluate detailed eye motions, specifically eyeball rotation jitter and pupil size jitter. We present our method to reduce the noise of motion captured eye movements without losing the saccadic motions. We found that the scaling values to use motion captured data vary depending on the virtual character and we computed values for the average jitter amplitudes of the eyeball rotation and the pupil size for three characters of different realism that can be used as a guideline. The average amplitude of the eye motion jitter perceived as most natural varies between 0.07 and 0.14 (see Table 1). The preferred average amplitudes of jitter affecting the pupil size range from 0.0011 and 0.0026 (see Table 2) in a scale where 0 would be a point size pupil and 1 would mean that the pupil fills the whole iris. These resulting values are very subtle but are still being perceived. Now that we have an indication of the levels of jitter perceived as most natural, we could create this noise procedurally. That way, one could build a tool for animators to enhance keyframed eye animations with jitter, leaving an adjustable parameter for scaling to adjust for the type of virtual character.

The two motions we used in our studies, a nine-point calibration motion and a reading motion, were both created to be neutral, showing no emotion and creating no arousal. However, for example, pupil unrest does vary depending on the person and situation. So it might be appropriate for animators to have a range of jitters to choose from even for the same character. It would be interesting to see if such small changes showing arousal or fatigue can be perceived by viewers.

Despite our post-processing technique, with our eye-tracker with its low frequency of 25Hz we were only able to detect saccades larger than about one degree in size. That means that it is very likely that some of the microsaccades in the signal were not detected. While we do not believe that this affected the average jitter in the angular rotations and pupil sizes that we found in a significant way – the noise perceived as most natural would still likely have a very similar average amplitude – it would be interesting to verify our results with more accurate data, ideally including microsaccades, ocular drift, and microtremors, and to test if this ground truth (signal alone without any noise) would be perceived as similarly or more natural compared to our results or if the addition of noise would still be necessary to yield the most realistic animations. Such data is difficult to obtain. While many eye-tracker models with higher framerates that allow for the detection of microsaccades exist, current commercially available technology still includes noise above the level that would be needed to correctly represent microtremors. But the technology is developing quickly and such data might soon become available. To determine parameters for animators to add noise to keyframed animations, it might be beneficial to not detect all microsaccades as the resulting detected signal might resemble more closely to a typical keyframed animation.

Finally, future work concerning eye motions also needs to consider periocular animation, e.g., of the lids and skin tissue surrounding the eye and needs to investigate smooth pursuit motions and higher-level eye movement behaviors such as eye movements during communication. The effect of vergence, unexpected events, or changes in light intensity would be further interesting avenues to explore.

ACKNOWLEDGEMENTS

This material is based in part upon work supported by the National Science Foundation under Grant No. IIS-1423189. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. We would like to thank Eakta Jain for her comments, and Adam Wentworth, Yuanchi Cai, Aubrey Lawson, Tyler Allen, Jordan Gestring, Galen Helfter, and Christian Sharpe for their contributions to modeling, capturing, and scripting.

REFERENCES

- [1] Ehud Ahissar, Amos Arieli, Moshe Fried, and Yoram Bonneh. 2016. On the possible roles of microsaccades and drifts in visual perception. *Vision Research* 118 (2016), 25–30. <https://doi.org/10.1016/j.visres.2014.12.004>
- [2] Sean Andrist, Tomislav Pejisa, Bilge Mutlu, and Michael Gleicher. 2012. Designing effective gaze mechanisms for virtual agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (*CHI '12*). 705–714. <https://doi.org/10.1145/2207676.2207777>
- [3] Pascal Bérard, Derek Bradley, Markus Gross, and Thabo Beeler. 2016. Lightweight eye capture using a parametric model. *ACM Transactions on Graphics* 35, 4, Article 117 (July 2016), 12 pages. <https://doi.org/10.1145/2897824.2925962>
- [4] Margaret M. Bradley, Laura Miccoli, Miguel A. Escrig, and Peter J. Lang. 2008. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 4 (July 2008), 602–607. <https://doi.org/10.1111/j.1469-8986.2008.00654.x>
- [5] Zhigang Deng, J. P. Lewis, and Ulrich Neumann. 2005. Automated eye motion using texture synthesis. *IEEE Computer Graphics Applications* 25, 2 (March 2005), 24–30. <https://doi.org/10.1109/MCG.2005.35>
- [6] Andrew Duchowski, Sophie Jörg, Aubrey Lawson, Takumi Bolte, Lech Świrski, and Krzysztof Krejtz. 2015. Eye movement synthesis with $1/f$ pink noise. In *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games* (Paris, France) (*MIG '15*). 47–56. <https://doi.org/10.1145/2822013.2822014>
- [7] Andrew T. Duchowski, Sophie Jörg, Tyler N. Allen, Ioannis Giannopoulos, and Krzysztof Krejtz. 2016. Eye movement synthesis. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (Charleston, SC) (*ETRA '16*). 147–154. <https://doi.org/10.1145/2857491.2857528>
- [8] Ralf Engbert. 2006. Microsaccades: a microcosm for research on oculomotor control, attention, and visual perception. In *Visual Perception Fundamentals of Vision: Low and Mid-Level Processes in Perception*, Susana Martinez-Conde, Stephen L. Macknik, Luis M. Martinez, Jose-Manuel Alonso, and Peter U. Tse (Eds.). Progress in Brain Research, Vol. 154. Elsevier B.V., 177–192. [https://doi.org/10.1016/S0079-6123\(06\)54009-9](https://doi.org/10.1016/S0079-6123(06)54009-9)
- [9] Ralf Engbert and Reinhold Kliegl. 2003. Microsaccades uncover the orientation of covert attention. *Vision Research* 43, 9 (2003), 1035–1045. [https://doi.org/10.1016/S0042-6989\(03\)00084-1](https://doi.org/10.1016/S0042-6989(03)00084-1)
- [10] Craig Evinger, Karen A. Manning, and Patrick A. Sibony. 1991. Eyelid movements. Mechanisms and normal data. *Investigative Ophthalmology & Visual Science* 32, 2 (1991), 387–400.
- [11] Maia Garau, Mel Slater, Simon Bee, and Martina Angela Sasse. 2001. The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (*CHI '01*). 309–316. <https://doi.org/10.1145/365024.365121>
- [12] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (*CHI '03*). 529–536. <https://doi.org/10.1145/642611.642703>
- [13] Jessica Hodgins, Sophie Jörg, Carol O'Sullivan, Sang Il Park, and Moshe Mahler. 2010. The saliency of anomalies in animated human characters. *ACM Transactions on Applied Perception* (TAP) 7, 4, Article 22 (July 2010), 14 pages. <https://doi.org/10.1145/1823738.1823740>
- [14] Jessica K. Hodgins, James F. O'Brien, and Jack Tumblin. 1998. Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics* 4, 4 (1998), 307–316. <https://doi.org/10.1109/2945.765325>
- [15] Birger Johansson and Christian Balkenius. 2017. A computational model of pupil dilation. *Connection Science* 30, 1 (January 2017), 5–19. <https://doi.org/10.1080/09540091.2016.1271401>
- [16] Sophie Jörg, Jessica Hodgins, and Carol O'Sullivan. 2010. The perception of finger motions. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (*APGV 2010*) (Los Angeles, California). 129–133. <https://doi.org/10.1145/1836248.1836273>
- [17] Krzysztof Krejtz, Andrew Duchowski, Heng Zhou, Sophie Jörg, and Anna Niedzielska. 2017. Perceptual evaluation of synthetic gaze jitter. *Computer Animation and Virtual Worlds* (2017), e1745. <https://doi.org/10.1002/cav.1745>
- [18] Mariska E. Kret, Karin Roelofs, Jeroen J. Stekelenburg, and Beatrice de Gelder. 2013. Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience* 7 (Dec 2013), 9 pages. <https://doi.org/10.3389/fnhum.2013.00810>
- [19] Sooha Park Lee, Jeremy B. Badler, and Norman I. Badler. 2002. Eyes alive. *ACM Transactions on Graphics* 21, 3 (July 2002), 637–644. <https://doi.org/10.1145/566654.566629>

- [20] Irene E. Loewenfeld. 1958. Mechanisms of reflex dilatation of the pupil. *Documenta Ophthalmologica* 12, 1 (01 Dec 1958), 185–448. <https://doi.org/10.1007/BF00913471>
- [21] C. E. Looser and T. Wheatley. 2010. The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science* 21, 12 (2010), 1854–62. <https://doi.org/10.1177/0956797610388044>
- [22] Otto Lowenstein, Richard Feinberg, and Irene E. Loewenfeld. 1963. Pupillary movements during acute and chronic fatigue: A new test for the objective evaluation of tiredness. *Washington, D.C.: Federal Aviation Agency, Office of Aviation Medicine, Reprinted from Investigative Ophthalmology* 2, 2 (1963), 138–157.
- [23] Xiaohan Ma and Zhigang Deng. 2009. Natural eye motion synthesis by modeling gaze-head coupling. In *IEEE Virtual Reality*. Lafayette, LA, 143–150. <https://doi.org/10.1109/VR.2009.4811014>
- [24] Karl F. MacDorman, Robert D. Green, Chin-Chang Ho, and Clinton T. Koch. 2009. Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior* 25, 3 (May 2009), 695–710. <https://doi.org/10.1016/j.chb.2008.12.026>
- [25] Susana Martinez-Conde, Stephen L. Macknik, and David H. Hubel. 2004. The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience* 5, 3 (March 2004), 229–240. <https://doi.org/10.1038/nrn1348>
- [26] Susana Martinez-Conde, Stephen L. Macknik, and David H. Troncoso, Xoana G. Hubel. 2009. Microsaccades: a neurophysiological analysis. *Trends in Neurosciences* 32, 9 (2009), 463–475. <https://doi.org/10.1016/j.tins.2009.05.006>
- [27] Rachel McDonnell, Martin Breidt, and Heinrich H. Bühlhoff. 2012. Render me real?: Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics* 31, 4, Article 91 (July 2012), 11 pages. <https://doi.org/10.1145/2185520.2185587>
- [28] Rachel McDonnell, Sophie Jörg, Jessica K. Hodgins, Fiona N. Newell, and Carol O’Sullivan. 2007. Virtual shapers & movers: Form and motion affect sex perception. In *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization (APGV '07)*. 7–10. <https://doi.org/10.1145/1272582.1272584>
- [29] J. W. McLaren, J. C. Erie, and R. F. Brubaker. 1992. Computerized analysis of pupillograms in studies of alertness. *Investigative ophthalmology & visual science* 33, 3 (March 1992), 671–6.
- [30] Aline Normoyle, Jeremy B. Badler, Teresa Fan, Norman I. Badler, Vinicius J. Cassol, and Soraia R. Musse. 2013. Evaluating perceived trust from procedurally animated gaze. In *Proceedings of the Motion in Games (Dublin 2, Ireland) (MIG '13)*. Article 119, 8 pages. <https://doi.org/10.1145/2522628.2522630>
- [31] Toshikazu Ohshima, Hiroyuki Yamamoto, and Hideyuki Tamura. 1996. Gaze-directed adaptive rendering for interacting with virtual space. In *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*. 103–110. <https://doi.org/10.1109/VRAIS.1996.490517>
- [32] Jonathan W. Peirce. 2008. Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics* 2 (2008), 8 pages. <https://doi.org/10.3389/neuro.11.010.2008>
- [33] Roy M. Pritchard. 1961. Stabilized images on the retina. *Scientific American* 204, 6 (June 1961), 72–79.
- [34] Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124, 3 (1998), 372–422.
- [35] Tobias Ritschel, Matthias Mittner, Jeppe Frisvad, Joris Coppens, Karol Myszkowski, and Hans-Peter Seidel. 2009. Temporal glare: Real-time dynamic simulation of the scattering in the human eye. *Computer Graphics Forum* 28, 2 (April 2009), 183–192. <https://doi.org/10.1111/j.1467-8659.2009.01357.x>
- [36] Kerstin Ruhland, S. Andrist, J. B. Badler, Christopher E. Peters, Norman I. Badler, Michael Gleicher, Bilge Mutlu, and Rachel McDonnell. 2014. Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. In *Eurographics 2014 - State of the Art Reports*, Sylvain Lefebvre and Michela Spagnuolo (Eds.). <https://doi.org/10.2312/egst.20141036>
- [37] Mark Sagar, Mike Seymour, and Annette Henderson. 2016. Creating connection with autonomous facial animation. *Communications of the ACM* 59, 12 (Dec. 2016), 82–91. <https://doi.org/10.1145/2950041>
- [38] Lawrence Stark, Fergus W. Campbell, and John Atwood. 1958. Pupil unrest: An example of noise in a biological servomechanism. *Nature* 182, 4639 (1958), 857–858.
- [39] William Steptoe, Oyewole Oyekoya, and Anthony Steed. 2010. Eyelid kinematics for virtual characters. *Computer Animation and Virtual Worlds* 21, 3–4 (May 2010), 161–171. <https://doi.org/10.1002/cav.v21:3/4>
- [40] Kazuki Takashima, Yasuko Omori, Yoshiharu Yoshimoto, Yuich Itoh, Yoshifumi Kitamura, and Fumio Kishino. 2008. Effects of avatar’s blinking animation on person impressions. In *Proceedings of Graphics Interface 2008 (Windsor, ON, Canada) (GI '08)*. Canadian Information Processing Society, Toronto, ON, Canada, 169–176. <http://dl.acm.org/citation.cfm?id=1375714.1375744>
- [41] Laura C. Trutoiu, Elizabeth J. Carter, Iain Matthews, and Jessica K. Hodgins. 2011. Modeling and animating eye blinks. *ACM Transactions on Applied Perception (TAP)* 2, 3, Article 17 (May 2011), 17 pages. <https://doi.org/10.1145/2010325.2010327>
- [42] Congyi Wang, Fuhao Shi, Shihong Xia, and Jinxiang Chai. 2016. Realtime 3D eye gaze animation using a single RGB camera. *ACM Transactions on Graphics* 35, 4, Article 118 (July 2016), 14 pages. <https://doi.org/10.1145/2897824.2925947>
- [43] Dong Wang, Fiona B. Mulvey, Jeff B. Pelz, and Kenneth Holmqvist. 2017. A study of artificial eyes for the measurement of precision in eye-trackers. *Behavior Research Methods* 49, 3 (01 Jun 2017), 947–959. <https://doi.org/10.3758/s13428-016-0755-8>
- [44] Sang Hoon Yeo, Martin Lesmana, Debanga R. Neog, and Dinesh K. Pai. 2012. Eyecatch: Simulating visuomotor coordination for object interception. *ACM Transactions on Graphics* 31, 4, Article 42 (July 2012), 10 pages. <https://doi.org/10.1145/2185520.2185538>

Received April 2018; revised May 2018; accepted June 2018