HAL
open science

Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems

Kerstin Ruhland, Sean Andrist, Jeremy Badler, Christopher Peters, Norman Badler, Michael Gleicher, Bilge Mutlu, Rachel Mcdonnell

To cite this version:

Kerstin Ruhland, Sean Andrist, Jeremy Badler, Christopher Peters, Norman Badler, et al.. Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. Eurographics 2014 - State of the Art Reports, Apr 2014, Strasbourg, France. pp.69-91, 10.2312/egst.20141036 . hal-01025241

HAL Id: hal-01025241

https://inria.hal.science/hal-01025241

Submitted on 17 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems

K. Ruhland[1], S. Andrist[2], J. B. Badler[3,4], C. E. Peters[5], N. I. Badler[4], M. Gleicher[2,6], B. Mutlu[2], R. McDonnell[1]

[1]Trinity College Dublin, Ireland
[2]University of Wisconsin, Madison, United States
[3]Northeastern University, Boston, United States
[4]University of Pennsylvania, Philadelphia, United States
[5]KTH Royal Institute of Technology, Stockholm, Sweden
[6]University of Grenoble and INRIA, France

Abstract

A person's emotions and state of mind are apparent in their face and eyes. As a Latin proverb states: "The face is the portrait of the mind; the eyes, its informers.". This presents a huge challenge for computer graphics researchers in the generation of artificial entities that aim to replicate the movement and appearance of the human eye, which is so important in human-human interactions. This State of the Art Report provides an overview of the efforts made on tackling this challenging task. As with many topics in Computer Graphics, a cross-disciplinary approach is required to fully understand the workings of the eye in the transmission of information to the user. We discuss the movement of the eyeballs, eyelids, and the head from a physiological perspective and how these movements can be modelled, rendered and animated in computer graphics applications. Further, we present recent research from psychology and sociology that seeks to understand higher level behaviours, such as attention and eye-gaze, during the expression of emotion or during conversation, and how they are synthesised in Computer Graphics and Robotics.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism - Animation—

1. Introduction

The generation of realistic artificial entities, from virtual characters to physical humanoid robots, has become a requirement underlying many applications in human-machine interactions. Apart from the obvious applications in the movie and special effects industries, they are needed for interactive applications such as games, virtual tourism and e-commerce, urban planning, medicine, surgery, training, and others. Realistic, plausible, and engaging characters contribute to high levels of immersion, enjoyment, and learning among human users. However, creating artificial entities which replicate humans in both appearance and behaviour remains one of the greatest challenges in the field of Computer Graphics. While their graphical fidelity has improved rapidly due to advances in graphics processing hardware and their accompanying rendering techniques, the qualitative mismatch between the appearance and behaviour of characters appears to be worsening. This can cause perceptual disparities that may disrupt viewers' sense of immersion, or may even contribute to unsettling emotional reactions [Mor70]. Behaviour synthesis has not yet reached the levels of realism required to convince users' perceptual systems that a virtual human is the real thing. This is particularly true when behaviour is generated automatically or semi-automatically in real-time, where a human animator or programmer has little or no intervention.

The human face is an important instrument for communicating and defining underlying emotions [EF03]. Eyes are central in conveying emotional information, and we are able to interpret the intentions and feelings of other humans

© The Eurographics Association 2014.

by observing their eyes. We learn this ability from a very young age, as infants have been shown to be efficient at distinguishing between emotional stimuli and other stimuli [Sim94]. The ability to use the eyes to guide and interpret social behaviour remains a central facet of social interactions throughout life, lending credence to the common proverb "the eyes are the window to the soul." It is not surprising, therefore, that much effort is required in the creation and animation of realistic virtual eyes.

Historically, implementations of eye gaze models typically have been proprietary in nature, often borrowing, somewhat arbitrarily, concepts from a variety of sources across a range of disciplines. One reason for this may be the difficulty of the endeavour, which requires the integration of knowledge from a large number of disciplines, such as psychology, neuroscience, and the social sciences. More recently, however, there has been an increase in the publication of comprehensive models, dedicated to virtual characters and robots, that integrate knowledge from these disciplines, making it possible to compile an overview of the process of designing gaze mechanisms appropriate for different applications. In order to design effective social gaze behaviours for robots and virtual agents, we must first gain a better understanding of low-level gaze cues (e.g., saccadic eye movements, mutual gaze, head motion, and so on), including the patterns in which they are produced and utilised in various social contexts. We also need to develop an understanding of how these low-level cues relate to high-level social and cognitive processes. From this understanding, we can develop computational models to synthesise gaze behaviour for agents and evaluate their effectiveness in eliciting positive outcomes among users.

In this survey, we present relevant and significant findings from anatomy and physiology on the construction and low-level workings of the human eye (Section 2 and 3). We also discuss research from the social sciences and psychology on how eye-gaze is used to convey information about direction of attention and emotional and mental states (Section 4). Throughout the report, we discuss how this information has been exploited in the development of eye-gaze models in computer graphics and human-robot interaction. We highlight some outstanding challenges regarding stylised characters and robots that result from their diverging proportions and handling (Section 5). Further, we report how researchers to date have evaluated the plausibility and realism of their eye models and conclude with future directions in the domain (Section 6 and 7).

We anticipate that this STAR will act as a focal point for knowledge in the computer graphics and animation community and will help to solidify research in this domain. The survey will be accessible to both established computer graphics researchers who wish to develop enhanced models of eye-gaze behaviours as well as to newcomers who wish to become familiarised with the most important sources in the domain.

2. Virtual Eye Creation

The synthesis of realistic artificial eyes requires the accurate modelling and rendering of eyeball structure. In this section, we discuss the basic anatomical structure of the human eye and present the most relevant research that uses this anatomical knowledge to produce geometry for facial animation or medical applications. The iris in particular presents a difficult challenge, as it is a layered and highly complex structure. Different approaches have been used to recover the iris structure and scattering features, from layering simple painted textures to recovering important details from photographs of an eye. This section focuses on the modelling and rendering approaches needed in order to achieve high realism in the structure and appearance of eye geometry.

2.1. Anatomical Structure

The basic anatomy of the human eye presented in this section is based on the book "Clinical Anatomy and Physiology of the Visual System" by Remington [Rem11]. The interested reader is referred to this book for more detailed information.

The eye is one of the most complex organs in the human body. Each individual layer of the eye performs a specific task, e.g. capturing light, focusing it and transforming it into electrical signals, to ensure correct perception of the environment. The transparent *cornea*, located at the front of the eye, is the first refraction point for light entering the eye (Figure 1). A tear film moistens and smoothens the surface of the

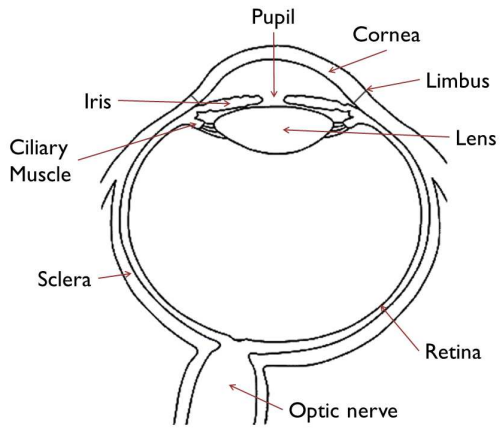

Figure 1: *Schematic of the Anatomy of the Human Eye*

cornea to ensure minimal scattering and distortion of light. Separated by the *corneal limbus*, the cornea is embedded in the outer shell, the so called *sclera*. Both can be pictured as two, not perfectly symmetric, spheres. The white opaque

© The Eurographics Association 2014.

tissue of the sclera preserves the eye's shape and provides protection against harmful influences and substances. The radius of curvature for the sclera and cornea varies between humans. On average, the sclera has a diameter of 12 mm and the cornea of 8 mm with a horizontal and vertical deviation due to their elliptical form.

The light passes through the visible *pupil* and is focused by the lens behind it. Depending on the prevailing light conditions, the muscles of the surrounding *iris* influence the shape and diameter of the pupil (Figure 2). The iris is a multilayered conical tissue. The pigmentation and density of the iris define eye colour, a unique identifier of a person. The iris' primary function is to regulate the amount of light that enters the inner eye. Light passing through the pupil is focused by the *lens* onto the retina. The lens' shape changes through contractions of the *ciliary muscle*, thereby increasing optical power and accommodating the projection of objects nearby onto the retina.

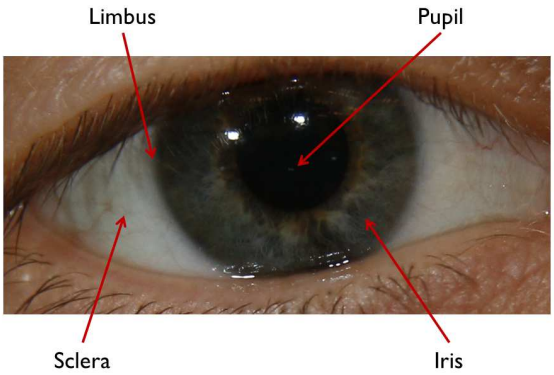

Figure 2: *Close-Up View of the Iris*

The *retina* forms the inner, light sensitive part of the eye. The light captured at the retina is processed and transmitted as electrical signals to the brain for further processing.

2.2. Modelling & Realistic Rendering

Eyes presented in animations or virtual environments are commonly modelled as spheres or half spheres and use high-resolution pictures of human eyes as texture maps [ID03, WLO10]. Only a few studies deal with a more anatomically realistic model of the eye, which are discussed in the following section.

For example, Sagar et al. [SBMH94] used simple Gouraud shading on two spheres, representing the sclera and cornea, and texture maps to model and render the human eye for a surgical simulation application. They realistically generated the retinal blood vessels shining through the sclera using a fractal tree [Opp86]. Two layers of polygons, with the inner layer oriented towards the pupil and the outer layer towards the corneal limbus, represented the iris fibres. When the pupil dilates in the model, Gaussian perturbation makes the outer layer wavy.

Other research has focused on replicating the visual appearance of the iris. Knowledge from ocularists was used to synthesise the human iris in Lefohn et al.'s [LBS*03] approach. Ocularists design aesthetically pleasing ocular prostheses by applying multiple layers of paint onto a plastic shell to paint the iris. Starting with the most dominant eye colour as a base, layers consisting of combinations of dots, radial smears, or radial spokes, representing iris components, are added. Lefohn et al. adapted this approach for a virtual eye by mapping scanned textures of painted layers onto overlaying frustums of right circular cones.

The most influential method for iris modelling to date is the image-based model of the eye by François et al. [FGBB09]. In an automated process, the iris morphology and light scattering features on the cornea and iris are extracted from iris photographs. The created iris subsurface map serves as render input for the subsurface texture mapping method. François et al. also take into account reflection and refraction at the corneal surface based on ambient light.

In contrast to the previously introduced image-based methods, the *ILIT* (Iridal Light Transport Model) by Lam and Baransoki [LB06] is an iris model based purely on studies from the fields of physics and biology. A Monte Carlo based rendering technique is implemented to replicate the interaction of the light within the iridial tissue and to determine the spectral response based on the known composition and structure of the iris and the properties of the light. The model allows the modification of biophysical parameters defining the composition of the iridal tissue and thereby affecting iris colour.

Other methods to synthesise iris images can be found in the field of iris recognition [WSG05, ZS06]. However, these methods are generated for verification of iris recognition algorithms and do not address the problem of modelling and rendering a realistic human iris.

An interesting area of research concerns the modelling of the pupil's reaction to the light. Pamplona et al. [POB09] presented a physiologically-based model for the pupil light reflex and implemented an image-based iris deformation model. In this model, a delay-differential equation was used to define the pupil diameter as a function of environment light, enhancing results from mathematical biology and experimental data. The authors' measurements of feature points on the iris during deformation showed that these points move along a straight line on radial axes. Therefore, the animation could be simplified to a texture-mapped planar triangle-strip mesh with the inner triangle vertices mapped to the pupil border. When changing the pupil diameter due to incident light, all vertices at the pupil's border are relocated to a new position along a radial line, connecting the vertex with the pupil's centre.

© The Eurographics Association 2014.

An additional reason for modelling pupil diameter is that it changes as a function of emotional arousal [BMEL08], and pupil dilation can function as an important social cue [TS04]. Obviously, emotion is expressed through eye gaze as well, which will be the subject of Section 4.3.

3. Low-Level Eye Animation

The first step in convincingly modelling any effective social gaze behaviour for virtual agents is to gain a thorough understanding of the underlying behaviour in humans. Fortunately, eye movements have been studied by psychologists and physiologists for more than a half-century, so an abundance of data exists that precisely describes the kinematics of eye movements.

Located in the frontal part of the human head, the eyes perceive only a restricted view of the surrounding environment. Humans' field of view is further limited due to the fact that high visual acuity is available only in a small region of the retina. Saccades, the vestibulo-ocular reflex, smooth pursuit movements and vergence (terms defined in later subsections) serve as mechanisms for the eyes to maintain fixation on a moving object, to stabilise the location of this object on the retinas and to adjust both eyes to visual depth. In addition, the eyes do not move in isolation: they are intricately intertwined with blinks and eyelid movements, and are very frequently just one component of a larger gaze shift that includes head motion.

In this section, the physiology underlying eye movements, eyelid movements and combined eye-head movements are reviewed briefly. We also highlight relevant articles from computer science and animation literature. Numerous approaches, including parametric approaches based on physiological measurements and computation approaches derived from motion capture or tracking data, have been used to successfully simulate realistic gaze in agents performing a variety of actions.

3.1. Eyeball Movement

The eyeball is a deceptively complex organ, capable of executing a wide repertoire of movements and controlled by six separate muscles (Figure 3) [LZ99]. Fortunately, these movements have been extensively studied by neurologists, psychologists and neuroscientists. As a result, both their characteristics and the conditions under which they occur have been extremely well-described, providing valuable resources for the animator who wishes to reproduce them.

3.1.1. Saccades

Saccades, if not the most frequent type of eye movement, are probably the most noticeable. They are the rapid shifts in eye position that centre the gaze on targets of interest. Saccade characteristics are quite stable across healthy individuals. Their duration increases approximately linearly and peak velocity increases exponentially with magnitude [Bec89], [LZ99, Chapter 2]. This consistent relationship is referred to as the main sequence (Figure 4) [BCS75].

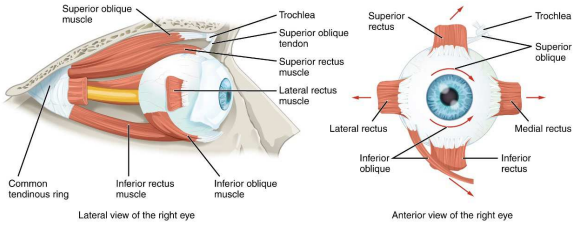

Figure 3: *Arrangement of the extraocular muscles. From "Anatomy & Physiology," Connexions, June 19, 2013, http://cnx.org/content/col11496/1.6/.*

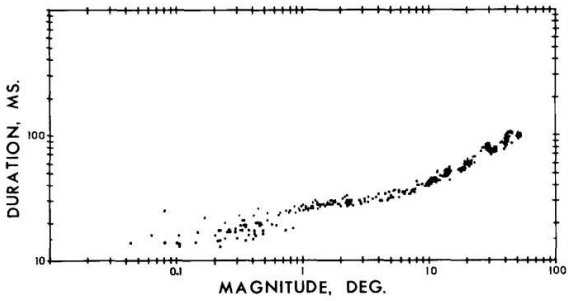

Figure 4: *Saccade duration versus magnitude over a wide range [BCS75, Figure 2]*

Initial acceleration and final deceleration are extremely rapid (> 10,000 deg/sec^2). For example, a very large saccade of 30 degrees typically has a velocity of around 500 deg/sec and a duration of less than one tenth of a second. Under natural conditions, such saccades are rare [BAS75]. More commonly occurring saccades of 5-10 degrees have durations of 30-40 msec, or approximately one frame at 30 Hz. Saccade latency in response to a visual target is normally around 200 msec [LZ99], but can reach approximately 100 msec under certain conditions [FR84]. Under most conditions, the smallest possible inter-saccadic interval (delay between consecutive saccades) is around 150 msec [LZ99]. Other saccade properties such as slight target undershoots, slow drifts after saccade completion (glissades), slight curvature in the spatial trajectory of oblique saccades, and torsional movements of the eyeball are all modest [LZ99] and are normally invisible to a casual observer.

Multiple approaches have been used to animate saccadic eye movements. Extremely faithful to the underlying muscle mechanics and innervation patterns driving the eyeball, Komogortsev et al. were able to generate extremely accu-

© The Eurographics Association 2014.

rate saccadic velocity curves [KHJK13]. Lee, Badler and Badler [LBB02] also attempted to faithfully reproduce a normalized velocity profile using a polynomial curve fit, and their results were used in some later work [GLBB07, QBM08, WLO10, Sha11, APMG12b]. Highly realistic movements can also be created using data-driven approaches; for example, by training a statistical model based on recorded movements [DLN05, DLN07, LMD12] or by simply replaying the records themselves (Figure 5) [HJO*10, MBB12]. An opposing perspective is exemplified by Yeo et al. [YLNP12], who implemented simple bell-shaped velocity curves with the rationale that subtle asymmetries of the velocity waveform were invisible to an observer. Similar methods use triangular velocity curves [VGSS04], uniform velocity [MH07], and instantaneous rotation [NBF*13] to model saccadic eye movements.

Figure 5: *Eye movement capture using a head-mounted eye-tracking device, and corresponding virtual avatar displaying the recorded eye-movements [MBB12]*

3.1.2. Vestibulo-ocular reflex

The Vestibulo-ocular reflex (VOR) acts to stabilise the eyes while they are fixating on an object during head motion. The VOR is modulated by a direct reflex via inner-ear vestibular neurons and thus occurs with extremely short latency, on the order of 7-15 msec [LZ99, Chapter 3]. It therefore can be considered as effectively simultaneous with head movement. It is easy to implement from a modelling perspective: if the head rotates with some angle θ_x, θ_y, θ_z the eyeballs should counter-roll at $-\theta_x$, $-\theta_y$, $-\theta_z$. If the eyeballs are radially uniform, their vergence angle is zero (see Section 3.1.4), and their θ_z component (corresponding to lateral head tilt) can be ignored as an additional simplification, see Anderson [And81] for an example of roll in a nonuniform (cat) eyeball. Note that the VOR is partially suppressed during large gaze shifts involving head movement (see Section 3.3). Most animation papers that parametrically simulated eye movements also included a VOR component [LBB02, GLBB07, MH07, WLO10, NBF*13, APMG12b].

3.1.3. Smooth pursuit

The human fixation system is not constrained to static objects. For example, it is able to make the complex series of saccades required to read an advertisement on the side of a moving bus. The smooth pursuit system is responsible for stabilising moving images on the retina [LZ99, Chapter 4]. It has a response latency intermediate to that of VOR and saccades, approximately 80-130 msec, and breaks down at high target speeds of 30 deg/sec or more [TL86]. Unlike VOR and saccades, smooth pursuit is more situational and therefore not often observed. It can be seen by watching spectators during a sporting match for example, or by watching a companion looking out the window of a moving train (technically this is optokinetic nystagmus, but it is visually similar to smooth pursuit). From a simulation perspective, pursuit is considerably harder to implement than saccades, since it requires a velocity calculation on top of the positional one. It is also harder to decide *what* to pursue, since motion needs to be added to any underlying attentional model (see e.g. Khullar et al. [CKB99], as well as Section 4). Due to its latency and limited velocity, it usually requires one or more catch-up saccades to be computed and executed, as well. For these reasons, it is currently of limited interest to animators, although the object-catching simulation of Yeo et al. [YLNP12] included an elegant implementation.

3.1.4. Vergence

Normally the two eyes are yoked; that is, if one eyeball rotates in a particular direction, the other will rotate in exactly the same direction. Vergence [LZ99, Chapter 8] is the exception to this rule. If an object lies on or near the visual midline, the two eyes must rotate in opposite directions to ensure that the object image appears at the corresponding position on both retinas. This process, called fusion, can accompany both saccades (e.g., if gaze changes from a far peripheral target to a close central one) or pursuit (if a moving target changes position in depth). Any animation system that calculates rotation angles for both eyes separately has *defacto* implemented vergence, but it is not strictly necessary for non-targeted gaze, such as that during speech (e.g., Masuko and Hoshino [MH07]) or emotional expression (e.g., Queiroz, Barros and Musse [QBM08]). Issues in vergence are exacerbated when the character has large or stylized eyes (see Section 5.1).

3.2. Eyelid Movement

The eyelids are not part of the oculomotor system, but they do interact with it. Their proximity to the eyeballs also necessitates a brief treatment here. Normal eye blinks can

© The Eurographics Association 2014.

be broken into spontaneous, voluntary, and reflexive subclasses, all of which have slightly different eyelid dynamics [VBR*03]. Spontaneous blinks, also called endogenous blinks, are the most interesting, since their frequency is linked to cognitive state and activity [SWG84, SNJ*07]. Various studies have linked blink occurrence to attentional processes [NKM*13], fatigue [JTC*07, AWH10], lying [BBW96], and speech production [NK10]. Blink rates are highly variable, however. A meta-study found ranges of 1.4-14.4/min during reading, 8.0-21.0/min during primary gaze and 10.5-32.5/min during conversation [Dou01].

The occurrence of individual blinks can be modelled as a Poisson process [Gre86]. However, blinks very often occur almost simultaneously with the onset of eye and eye-head gaze movements, particularly large ones over 30 degrees [EMP*94]. It is also important to note that during a single blink, eyelid motion is *not* uniform: the down-phase velocity is approximately twice as fast as that of the up-phase velocity, and their respective durations are nonlinear (Figure 6) [EMS91, GSC91]. Importantly, eyelid displacements always accompany vertical saccades called lid saccades. Lid saccades do not exhibit as much marked asymmetry between down- and up-phases as do blinks (Figure 7) [EMS91, GSC91].

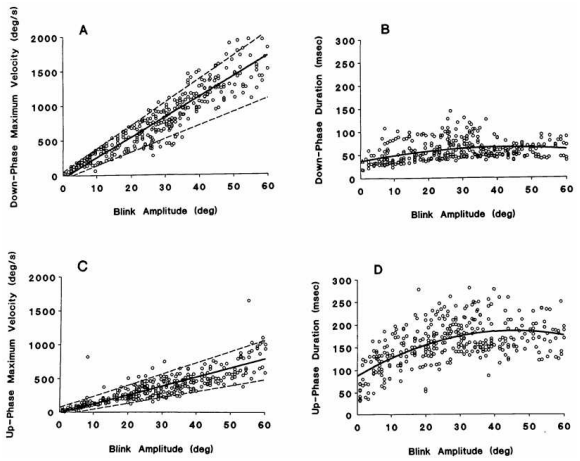

Figure 6: *Different eyelid velocities and movement durations for the down- and up-phases of a blink [EMS91, Figure 2]*

An unblinking face is visually disconcerting enough that nearly all graphics papers implement some form of blink production; however, the blinks are usually not described in detail, leaving their precise implementation to be handled by the facial animation package used. It is possible to animate blinks by calculating eyelid trajectories based on recorded video [DLN05, DLN07, WLO10, LMD12]. More sophisticated blink models also exist [SOS10, TCMH11], in which eyelid dynamics are modelled according to physiological data and account for both endogenous blinks and eyelid saccades. Normoyle and colleagues [NBF*13] used a heuristic approach based on physiology: eyelid displacement was proportional to eyeball rotation except for the downward blink phase, which was modelled as a step displacement. A number of eye blink model focus on higher-level aspects, such as the timing and synchronisation of blinks during head movements and conversations [LBB02, GLBB07, MH07]. Peters [Pet10] compared the realism of different methods for simulating blink timing, such as constant frequency versus saccade-linked.

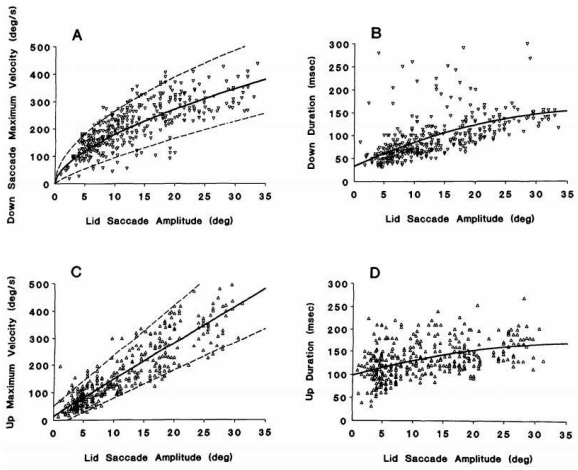

Figure 7: *Different eyelid velocities and movement durations for the down- and up-phases of a blink [EMS91, Figure 7]*

3.3. Combined Eye-Head Movement

Most natural gaze shifts employ both eye and head movements to replicate human gaze behaviours (Figure 8) [LZ99, Chapter 7]. The threshold for evoking an eye-head gaze shift as opposed to an eye-only saccade is approximately 15-20 degrees [Sta99], although this varies between individuals [Ful92]. When reacting to a target, eyes normally move first with typical saccadic latency (~200 msec), while head motion begins 20-50 msec later. However, when gaze is shifted to a predictable target, the head movement begins around 100 msec before the eyes [LZ99]. Like eye saccades, head movements have a consistent velocity-magnitude relationship [GV87]. Unlike saccades, subjects can voluntarily modulate head velocity with a corresponding impact on eye velocity. (Figure 8B) [LR86].

A common strategy employed in graphics literature is to define a cutoff displacement of 10-15 degrees above which targets are acquired by an eye-head gaze. Once the desired head displacement is known, there are several possibilities for computing the rest of its movement param-

© The Eurographics Association 2014.

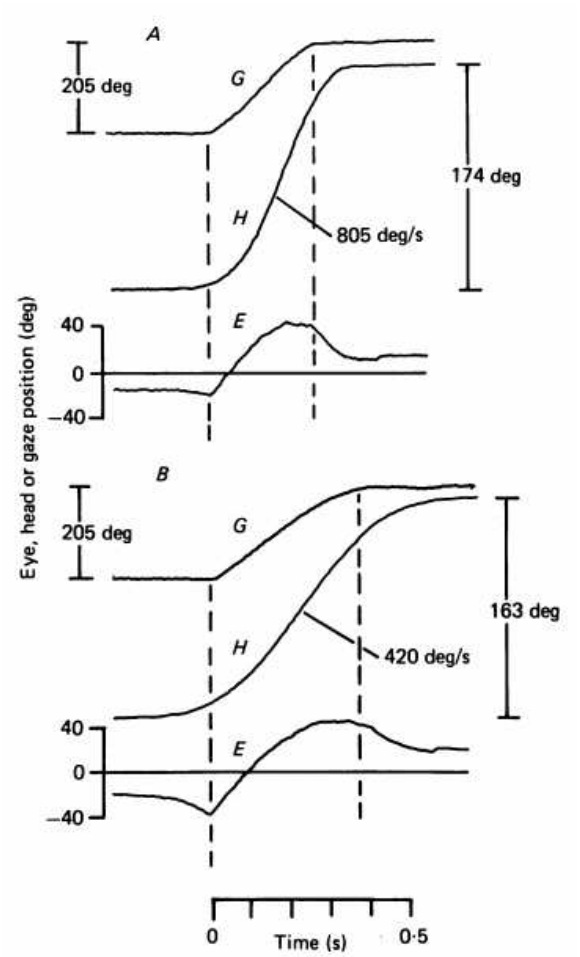

Figure 8: *Combined eye-head gaze shifts. E signifies eye, H head and G gaze [LR86, Figure 1]*

eters. Its velocity can be estimated using the main sequence relation [GV87, YLNP12] or from a desired timing regime [MH07], or simply computed automatically using inverse kinematics of the head and neck joints [NBF*13]. Head animation can also be modelled by fitting physiological [Pet10] or motion capture [VGSS04, LMD12] data. The model of Andrist et al. [APMG12b] attempts to provide parametric control over many of the aspects of movements that vary based on actor and situation while maintaining the low-level kinematics common to all eye-head movements.

3.4. Summary

Eye movements, blinks, and head gaze have all been extensively studied by physiologists, so their parameters are quite well understood. As a result, animators have access to a wealth of data that can be used to increase the realism of their own simulations. An important consideration is to decide which details are important for imparting realism, and which details are too subtle or unnoticeable to be worth incorporating into virtual agents.

Changes to low-level features of gaze motion, such as direction, velocity, and amount of head recruitment, can produce substantial effects on high-level outcomes of the interaction. For this reason, it is important to parameterise the low-level features in such a way that designers of human-virtual agent interactions can easily target specific high-level outcomes through the manipulation of low-level parameters. For example, Andrist et al. [APMG12a] have shown that manipulation of the amount of head alignment in a gaze shift toward users and objects can serve to create gaze behaviour that is perceived as more affiliative–leading to higher feelings of rapport–or more referential–leading to learning outcomes–in an educational scenario.

In this section, a great deal of attention was paid to movement kinematics. However, there are other characteristics of eye movements that remain to be resolved; namely, when they occur, where they are directed, and the cognitive state of the agent executing the movement. These higher-level issues will be addressed in the following section.

4. High-Level Aspects of Gaze Behaviour

This section is founded on the low-level aspects of combined eye-head animation described in Section 3. It considers the control of eye-gaze in virtual characters for the purposes of allowing them to investigate their environments, to convey emotional expressions, and to engage in social interactions.

We discuss the use of computational visual attention models to direct the eye-gaze of virtual characters to areas of relevance and potential interest in virtual environments. Visual attention models are discussed to motivate the provision of target locations to drive eye-gaze motions and the expressive quality of those motions. As such, dwell duration, blinking, and eye-head ratio may convey varying levels of interest a character has in an event, for example.

We also review the latest research concerning eye-gaze synthesis and control for creating embodied conversational characters capable of engaging in natural interactions with human users in real time. This will cover approaches which involve copying human eye-gaze behaviours and low-level eye-gaze parameters, such as the durations and timings of mutual gaze and gaze aversions, onto virtual characters, for the purpose of conveying expressions of interest, affection, and dominance. Multi-modal issues are important here, since synchrony between eye-gaze and other expressive modalities is vital to ensure fluid communication and turn-taking. Variations of eye-gaze with respect to conversational roles of speaker and listener are also discussed in relation to the creation of characters with specialised behaviours. These include, for example, patterns of eye and gaze behaviour occurring during specific speech acts and 'listening agents' ca-

© The Eurographics Association 2014.

pable of providing real-time non-verbal (e.g. eye-gaze, head-nods) and paraverbal (e.g. verbalisations such as "Uh-huh") feedback to users to maintain conversation engagement.

4.1. From Humans to Characters

Gaze serves a number of functions, including information gathering, signalling of interest and emotional state, and regulating conversations through management of participant turn-taking, that are well-surveyed in the literature [AC76, Ken90]. A common purpose of gaze models in computer graphics and animation is to provide a window into the mind of characters [LMT*07]; that is, to communicate information to a viewer about the internal states, attitudes, attentions, and intentions of characters [MGR04]. Gaze behaviour is also crucial for autonomous virtual agents that utilise it as a means of mobile sensory investigation and for characters whose primary role involves real-time interactions with human users [Cas00].

A key challenge is to automate a process often conducted manually by artists. Models must be capable of fully automating expressions appropriately given the context of the situation (i.e. environment, interaction type, culture, and so on). In this respect, automation efforts for characters [CV99] share many cross-overs with efforts made in social robotics (Section 5.2), especially when characters are to be endowed with behaviour generation competencies that are associated with behavioural and cognitive animation approaches in computer animation [TT94, FTT99]. Research efforts animate gaze for virtual characters across a range scenarios, including walking down (virtual) streets, engaging in communicative activities with others (humans and agents), and behaving in a human-like manner even in the absence of a particular task.

In this section, competencies are considered in areas of generating target locations in scenes (Section 4.2); how eye-gaze looking behaviours should be generated in order to properly express emotions (Section 4.3); how gaze may open and regulate interaction, signal engagement, and act as attentive listener (Section 4.4); and how gaze may be used with other modalities, including speech, to enable characters to more naturally engage in discourse (Section 4.5).

4.2. Visual Attention

Automating eye and gaze behaviour relates to the question of target location [CKB99]. Many research efforts have focused on the use of visual attention models to control and direct characters' gaze toward locations of importance in both virtual and real environments. These models are important so that characters are perceived as dynamically responsive to events occurring in their environment, and these models are useful for the generation of autonomous secondary behaviour [GD02].

Gaze behaviour is also crucial for autonomous virtual agents that utilise it as a means of mobile sensory investigation. Visual attention may be coupled with visual perception input systems (see Peters et al. [PCR*11] for an overview of synthetic vision and casting ray approaches) in order to allow characters to query their virtual surroundings. Based on perceptual inputs, visual attention models determine locations of interest to which gaze and eye movements are subsequently directed. These approaches have been employed in the animation of idle looking behaviours of characters in a bottom-up manner, i.e., when there are no tasks at hand or to interrupt tasks. They have also been applied to situations in which the character is embedded in virtual environments [POS03]. More recent work [CGV09] considers relationships between gaze targets and proxemics.

In other situations, the virtual character interacts with the real environment through a web camera [ID03, PBER07] and makes eye movements and gaze motions toward salient locations in the environment. Such saliency-based approaches [IDP06, OSS09] are based on a neurobiological model of visual attention [Itt00] and have been popular for animating eye and gaze movements. More recent efforts have focused on specific aspects of visual attention; for example, the role of object relevance [OSP11] and task constraints [KOS11] on gaze specification. A key challenge is balancing bottom-up and top-down visual attention for gaze allocation [MHKS07]. This is a difficult issue due to the demands of real-time operation.

4.3. Expression of Emotion

Gaze is a powerful method by which emotions are expressed [Iza91]. Therefore, how characters perform gaze motions is an important consideration for successful conveyance of emotional states to users [FOM*02].

While traditional approaches by skilled animators have been successful at creating expressive characters, automatic approaches are still an open challenge. Current approaches in the literature have therefore sought to analyse gaze motions in animated films to create animation models that can automatically map between emotions and gaze animation characteristics [LMK04, QBM07, LM10a].

Studies have also considered the expression of emotions through gaze shifts that involve movements of both the torso and the head [LM07]. For example, the SmartBody Gaze Controller [TLM09], released as part of the open-source SmartBody system, is capable of making full-body torso gaze shifts and aversions and also allows for the control of parameters, such as velocity and postural bias. Empirical studies have been performed in order to link low-level gaze attributes from non-verbal behaviour literature with observers' attributions of emotional states [LM10b, QBM08]. For example, Cig et al. [CKEMT10] conducted user studies to show that changes in gaze and head behaviour, achieved through the variation of gaze animation parameters, led to

© The Eurographics Association 2014.

changes in impressions of the arousal and dominance levels of characters. Li et al. [LM12] describe a rule-based approach to generate emotional eye movement based on the Geneva Emotion Wheel to enable virtual agents to convey different emotional expressions to users through eye movements. A data-driven approach was adopted by Busso et al. [BDNN07] to generate expressive head movements from speech data.

4.4. Non-Verbal Interaction

Research in psychology has revealed gaze to be a powerful non-verbal cue that yields a number of positive effects in human-human interactions [Kle86]. Gazing at transition-relevant places in a conversation can facilitate conversational turn-taking by reducing the length of pauses between speaker turns and reducing the amount of overlapping speech [Ken67], and a teacher who engages in mutual gaze with students is effective in improving learning outcomes [OO80]. Gaze is used to regulate intimacy [Abe86], facilitate references to objects in a shared visual space [BHAF09, GC11], signal dominance in multiparty conversations [FN11], and facilitate conversational grounding, which leads to improved collaborative outcomes [BHZS12]. People take another's eye gaze as evidence of what they are attending to and thinking about. They then use that evidence in determining the course of their current utterance or action [CK03]. Eye gaze produced by a speaker can be used by an addressee to resolve conversational ambiguity [HB07]. When multiple interactions take place over a long period of time, patterns of gaze and other non-verbal behaviours have been shown to adapt as relationships evolve [SB12].

By deploying an agent's gaze strategically, a number of the above-described positive high-level outcomes can be achieved in conversations between humans and agents. Wang and Gratch [WG10] found that a virtual agent exhibiting gaze attention and displaying cues of positivity and coordination to a user can create stronger feelings of rapport from the user. In an interactive storytelling scenario, a virtual agent that modulates mutual gaze by shifting its gaze in reaction to a user's gaze is able to improve user perceptions of social presence and rapport [BWA*10]. In immersive virtual environments, a virtual agent has been shown to influence the amount of interpersonal distance a human user will maintain with the agent based on the amount of eye contact the agent makes with the user [BBBL01]. Appropriately-timed shifts in gaze away from the user–in relation to the speech of the user–can lead to positive high-level outcomes in turn-taking and participant disclosure [AMG13]. An agent can also use gaze (along with gesture and speech) to effectively shape a conversation with multiple participants according to its own intentions [BH10].

4.4.1. Conversation and Listening Behaviour

Non-verbal feedback behaviours relating to gaze, such as glances towards and away from others, and other modalities have many functions [CTP98] in mediating flow in conversational situations [CT99], including indicating the addressee, paying attention, displaying attentiveness, affecting turn transitions, and signalling requests for backchannels [Hey06]. Previous approaches to modelling conversational gaze for characters include those considering communicative functions and statistical information of gaze patterns to generate directed and averted gaze for dyadic interactions [PB03]. Other approaches have used visual attention models to simulate social gaze, with accompanying engagement and distraction behaviours, during multiparty interactions [GB06, GLBB07].

The effects of cultural differences and gender on the conversational behaviour of agents is also of importance. Jan et al. [JHM*07] have simulated different cultural parameters related to gaze, overlap in turn-taking, and proxemics for rating by native speakers of North American English, Mexican Spanish, and Arabic. Studies in immersive and augmented reality environments have shown that users provide more personal space to agents that engage in mutual gaze with them [BBLB02], and that users have a higher physiological arousal toward virtual agents not exhibiting behaviours of their cultural background [ODK*12]. In Vala et al. [VBP11], gaze is considered in the creation of a model for varying the communication of agents based on gender.

Recent efforts have involved the use of comprehensive annotated databases of conversations between humans and characters, for example [MVC*12]. Approaches also consider the animation of gaze aversion behaviours for virtual agents in conversation situations [AMG13].

A notable area of research attention involves the definition of non-verbal signals for artificial listeners [HNP07] (see Bevacqua [Bev13] for an overview). In these situations, gaze has been studied in the context of backchannels during conversation, i.e., non-intrusive acoustic and visual signals provided to the speaker by listeners during their turn [YNG70]. The Rapport Agent, developed by Gratch et al. [GWG*07], provides non-verbal backchannels for the listener. More recently, Hjalmarsson and Oertel [HO12] have found that listeners are more prone to backchannel when the gaze of a virtual agent is directed towards them; however, gaze alone cannot explain the timing of backchannels.

4.4.2. Opening interactions and directing attention

The ability of characters to express interest [GP09] in users through gaze [PPB*05] and associated modalities has led to research on less explicit and more natural ways in which interaction with humans is requested and human attention is directed. The ability to express interest through eye movements and gaze has been studied in relation to its effects

© The Eurographics Association 2014.

on human perception of attention [Pet06] and human self-involvement during social interactions [MSH*06] with characters. Other studies have considered the use of the eyes, head, body, and locomotive direction of characters [Pet05] (see Figure 9) as a means for opening an interaction [Car78].

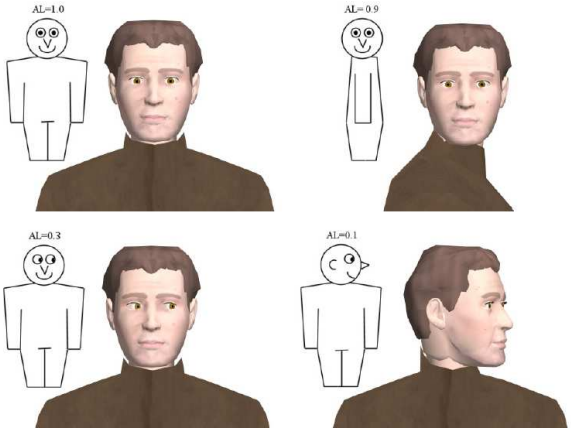

Figure 9: *In [Pet05] the eye, head, and body directions of a character during gaze behaviours are associated with perceived attention levels for evaluating conversation opening [Pet06] at varying distances in virtual environments.*

Such an approach is useful in situations where verbal utterances may not be desirable or practical, due to communication distance or for fear of negative social repercussions arising from failed attempts [Gof63]. Other work has focused on characters that initiate interactions in a favourable manner using gaze in order to promote subsequent interaction attempts [BAT09]. Cafaro et al. [CVB*12] further investigate how the multimodal non-verbal behaviours of characters during the approach to interaction leads to the formation of impressions related to personality and interpersonal attitude. They also highlight the relationships between gaze and judgements of friendliness.

Gaze behaviours may be a powerful means for cueing the attention of human users [FBT07]. A number of studies have investigated the gaze cueing abilities of virtual characters [MSSSB10] in game scenarios [PAK09] and in multi-task scenarios [KK13]. Further research is required to elucidate the factors that may underlie the ability of artificial systems to better direct the visual attention of human users. This offers significant potential to improve human task performance and provides a basis for implementing joint attention capabilities between humans and characters.

4.5. Verbal Interaction

Gaze is an important component in multi-modal behaviours conducted by speakers in face-to-face interactions, and is only one of a number of non-verbal channels typically employed in parallel to speech to efficiently request, take, hold, and give interaction turns [Dun74] and request feedback from listeners [Cho92]. The generation of coordinated verbal and non-verbal multimodal behaviours, including gaze [PPDR00], has therefore been an important focus of research for virtual characters engaging in conversation with humans. Sometimes this involves artificial systems being able to detect engagement through gaze from real conversational partners [NI10, IONN13] in order to adapt to them.

4.5.1. Speech driven gaze

A number of systems use speech as an input from which to generate facial expressions involving the mouth, head, eyes, and eyebrows [AHS02]. More recently, Zoric et al. [ZFP11] automatically generated facial gestures in real-time from the prosodic information obtained from speech signals. Nods, head movements, blinks, eyebrow gestures, and gaze were generated using Hidden Markov Models and global statistics. Gaze level fell at the hesitation pause, and rose at the end of the utterance in order to obtain listener feedback. Le et al. [LMD12] generated head motion, eye gaze, and eyelid motion simultaneously from speech inputs. Non-linear dynamical canonical correlation analysis was used to synthesise gaze from head motion and speech features. Mariooryad and Busso [MB12] focused on the generation of head and eyebrow motions from speech using three Dynamic Bayesian Networks (DBNs). In recent work, Marsella et al. [MXL*13] utilised semantics based on a shallow analysis of the utterance text, and prosody from the speech signal, to generate head movements, eye saccades, gestures, blinks, and gaze behaviour. Their method outperforms previous approaches that used prosody alone.

4.5.2. Gaze and verbal behaviour

A number of research efforts have considered eye movements and gaze specifically during situations involving verbal communication between characters and humans [CCD00]. Vinayagamoorthy et al. [VGSS04] presented an eye-gaze model for user-controlled avatars involved in dyadic interactions in shared immersive virtual environments. Their model accounts for differences in saccade magnitude and the effect of the roles of listener or speaker on gaze and inter-saccadic interval, and also generates movements for other parts of the body. Breton et al. [BPG*06] modelled the gaze behaviour of conversational characters in real-time multimodal dialogue with a group of users. Ishii et al. [IMFN06] proposed a model for animating the gaze behaviour of an avatar for turn-taking in multiparty conversations using utterances and a probabilistic state transition model.

© The Eurographics Association 2014.

4.5.3. Gaze and multimodal behaviour generation

The automatic generation of conversational behaviour for characters [CPB*94] involves the generation and coordination of verbal and non-verbal behaviour over multiple modalities, such as facial expressions, lip motions, eye gaze, head motion, and arm gestures in order to create meaningful gestures (Figure 10). Systems, such as BodyChat [VC98],

Figure 10: *Gaze as one of a number of coordinated modalities expressing panic and fear in an Embodied Conversational Agent (ECA) [NHP13]*

BEAT [CVB01], and Spark [Vil04], allowed characters to automatically animate their own conversational behaviours, including speech, gaze, turn-taking, and backchannel feedback, with minimal user and animator specification through text input, for example. Bee et al. [BPAW10] combined gaze and linguistic behaviours to create agents capable of expressing social dominance. For an overview of multimodal coordination, see Martin et al. [MDR*11].

Initiatives such as SAIBA (Situation, Agent, Intention, Behaviour, Animation) [KKM*06] are being developed to modularise the design of conversational characters. Behaviour Markup Language (BML) [VCC*07], developed as one of three stages in the SAIBA model, defines multimodal behaviours, including gaze, head, face, body, gesture, speech, and others, in a human-readable XML mark-up format. BML allows the definition of multimodal behaviours by specifying temporal details for primitive action elements (see Krenn et al. [KPPP11] for an overview).

5. Unique Challenges

A unique challenge facing the modelling of eye gaze behaviours for graphics applications is that not all characters conform to human anatomical proportions. This can lead to undesirable artefacts, such as cross-eyedness in characters with non-human or exaggerated human geometry. Existing algorithms must retarget animation and alteration in order to produce plausible results.

We also discuss challenges in robotics that may have interesting crossovers with efforts in the domain of virtual characters. In addition, we describe the mapping of gaze behaviours between virtual and physical embodiments. A number of challenges, discussed in Section 5.2, arise in the retargeting of behaviours from a virtual agent to a humanlike robotic platform. One of these challenges for example, is the attempt to retarget social gaze behaviours from a virtual platform in which agents can use both their head and eyes to a robotic platform without articulated eyes. The investigation of reduced embodiments is important to computer graphics applications where lower fidelity or more constrained characters are required. Nonetheless, this attempt must ensure to convey similar communicative qualities to their high fidelity counterparts.

5.1. Stylised Characters

Most gaze models are based on observations of humans or other animals that have eyes that are of a particular size and movements that are subject to physical properties and anatomical limits. Real-world gaze movement must be functional: the person or animal must be able to see what they need to see. When gaze models are applied to characters, such as cartoon or stylised characters, whose eyes are significantly different, problems can occur. The most basic question is whether or not it is appropriate to use models developed for humans and human eyes for characters. Lacking an alternative, this is the most common approach. The human brain has a particular affinity for faces and eyes (e.g., Bentin et al. [BAP*96]), and peoples' propensity to interpret even vaguely eye-like objects as eyes suggests adapting human models to characters is a viable approach.

In applying human gaze models to characters with eye geometry different from that of realistic humans (see Figure 11), a number of issues arise. Pejsa et al. [PMG13] catalogue and propose solutions to several of these issues. The issues fall into three key areas. The first is that "correct" human movements often look odd when examined closely. Human and animal eyes are sufficiently small that most people never see the details of their movements. When these movements are magnified, for example by using a character with proportionally larger eyes, the movements are unexpected. Second, stylised character eyes are often larger, and therefore would require unrealistically fast movements to traverse the angular displacements of real eye movements. Third, stylised eyes often have geometric configurations not seen in human eyes, such as asymmetric shapes and movement ranges. Each of these issues leads to similar effects: when human eye models are applied to stylised eyes, visual artifacts occur, such as movements that seem unusual or would not be seen in humans except under pathological circumstances (such as cross-eyedness or divergence).

To adapt a human gaze model to a range of stylised characters (see Figure 11), Pejsa et al. [PMG13] added constraints to ensure that pathological situations, such as cross-eyedness, do not occur. To meet these constraints, they exploited the fact that virtual agents are performative. They do not need to function correctly (e.g. enable the character to see); rather, they serve as a signal to the audience. The approach moves gaze targets and limits eye movements in or-

© The Eurographics Association 2014.

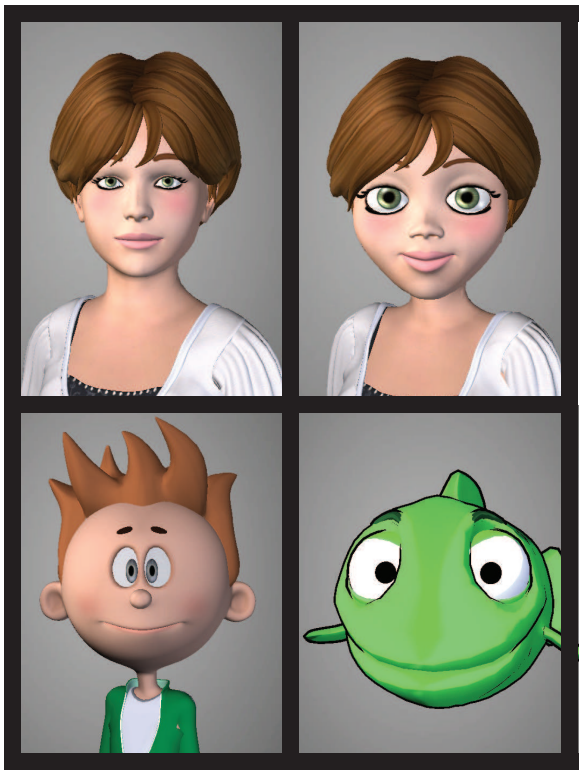

Figure 11: *Four different characters with different eye geometry. The character on the upper left has realistic human proportions. Adapting gaze models to characters with stylised eye geometry requires considering potential artifacts. [PMG13]*

der to avoid objectionable artifacts. While the model did not accurately point the eyes at the intended gaze targets, an empirical study confirmed that this inaccuracy did not damage viewers' ability to localise gaze targets and that the visible artifacts were significantly reduced.

5.2. Robots

The virtual agents and social robotics communities have separately investigated a number of gaze mechanisms and their conversational effects. However, it is often unclear how these mechanisms might be accurately translated between the virtual and physical realms. Robots and virtual agents have been shown to differ along a number of social dimensions, including realism, social presence, lifelikeness, and physical proximity [PKFT07]. A number of studies have been performed that compare virtual agents with humanlike robots in various contexts, such as health education and collaborative work [PKFT07, KB04, BHKS11].

A number of unique challenges arise when attempting to repurpose behaviours from a virtual agent for a humanlike robotic platform, including (1) the acceleration and speed of a robot's movements have both upper and lower limits, (2) due to physical inertia and communication latency, a robot will typically not react instantaneously to a command, and (3) robot expression has fewer degrees of freedom [LvW12]. The latter challenge is of particular importance when attempting to bring gaze mechanisms from a virtual platform, in which agents can use both their head and eyes, to a robotic platform without articulated eyes. In a similar vein, the physically co-located robot must actively direct its gaze to track users' faces, motions that are not required of a virtual agent due to the Mona Lisa gaze effect [DdGB10, MEB12]. Conversely, this may also enable physical systems to better direct human gaze to targets of interest in the real environment. In general, retooling a behaviour for a completely different platform (e.g., translating a virtual agent gaze model for humanlike robots) requires a consideration of what minimum representation that behaviour would need in order to evoke the desired response [CT07]. Recent work by Andrist et al. [ATGM14] showed how conversational gaze aversion behaviours, originally developed for virtual agents, could be effectively adapted for a humanlike robot platform that lacks articulated eyes (Figure 12).

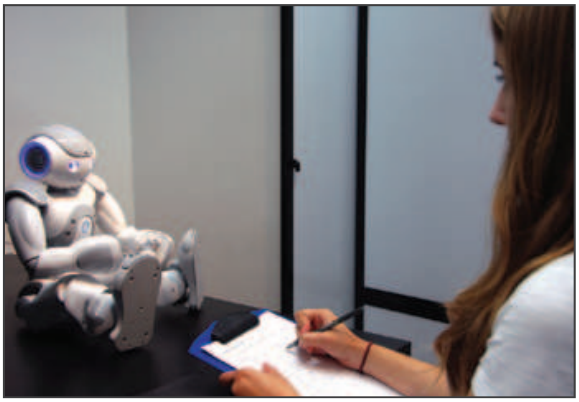

Figure 12: *A human conversational partner interacting with the NAO robot [ATGM14]*

In the human-robot interaction (HRI) community, a particular emphasis has been placed on examining the high-level outcomes achievable through gaze. This focus may be due to the HRI community's strong ties to social science communities, and partially due to the fact that HRI researchers generally have less control over the low-level variables of gaze. For example, HRI researchers are limited by the velocities achievable by physical motors, and affordances of their robots or when working with a robot that does not have articulated eyes and must express gaze through head movements alone. Srinivasan and Murphy [SM11] provide a survey of HRI gaze research that identifies a number of social contexts and discrete gaze acts considered so far by the community.

© The Eurographics Association 2014.

A number of high-level conversational outcomes achievable through gaze behaviour have been investigated in HRI research, including the use of gaze to signal participant roles [MSK*09], facilitate turn-exchanges [KYH*98], and affect the human user's physical and psychological closeness with the robot [MM11]. A robot that gazes responsively toward the human user, rather than randomly or statically, produces stronger feelings of being intentionally looked at [YSI*06]. Previous research has also studied how to deploy a robot's gaze as a cue for its attention, based on detected engagement levels of human participants in the interaction [XLW13], and timed in relation to object references made by the robot when speaking or by the user when the robot is listening [SC09]. In conversations with multiple participants, a robot's scanning gaze behaviour serves the dual functions of indicating attention to all conversational partners and updating the robot's knowledge of the partners that are occasionally outside the range of its cameras [BFJ*05].

6. Evaluation

Researchers strive to create virtual entities that evoke the same reactions as real humans do during social interaction. It is therefore important to take human perceptions into account in judging the plausibility and realism of virtual entity simulations. In this section, we discuss the ways in which low- and high-level gaze models have been evaluated in terms of plausibility, effectiveness, and ability to achieve communicative goals. Various approaches have been adopted, including survey-based questionnaires, direct comparisons with state of the art algorithms, and ground truth eye data captured from a human, to achieve this goal. We highlight the importance of perception and how effective perceptual experiments can be designed.

6.1. Experimentation methods

The effectiveness of an experiment depends on the careful preparation of stimuli, the presentation of those stimuli, and a correct analysis of the results. This section describes in detail these individual steps.

6.1.1. Stimuli Preparation

The development of an experiment begins with the acquisition of stimuli that will be shown to the participants. These stimuli are typically images or animations which focus on a specific factor of interest. There are a number of different media and methods for presenting stimuli, such as videos, interactive applications or virtual reality. In the case of evaluating the accuracy of a head-eye coordination model, the stimulus presented might be an animated character gazing toward objects on a desk, as found by Andrist et al. [APMG12a]. In this study, a series of videos was generated, and one of two videos were shown to participants. In one video, an animated virtual character performed gaze-shifts generated by the proposed model. In the other video, the state-of-the-art model and a human confederate looked at objects. Sixteen objects were placed in two rows on a desk with form or colour coding for easy identification. Communicative accuracy was measured based on how accurately the participants identified the object towards which the agent was looking. In work focusing on eye-representation, Steptoe and Steed generated a scene in which an avatar was seated behind a clock-face and viewed from different camera angles [SS08]. Only the avatar's eyes were directed toward a specific number on the clock-face. Conditions alternated with respect to eye socket deformation and vergence. The authors showed that realistic socket motion increased both perceived authenticity and the viewer's self-assessment of having correctly identified the point of regard, while vergence had a much more modest effect.

In order to investigate if a proposed model represents an improvement on previous methods, comparisons may be performed between sample stimuli. For example, Steptoe et al. [SOS10] generated a series of videos which animated blinks and lid saccades of two different virtual characters based on motion captured data, their proposed models, and simple linear interpolation. Animations of each condition assembled in one video were then ranked regarding perceived realism and similarity to the recorded video. Similarly, Deng et al. [DLN05] synthesised eye motions with different existing approaches and compared the results with their own model. This allowed them to test if their model showed an improvement over previous methods in naturalness. Peters et al. [PQ10] focused on identifying the most appropriate parameters of an eye gaze and blinking system and created animations which altered only the blinking conditions, eye-head ratio, or direction of head movements to assess which strategy would be better received by users. With a similar aim, Trutoiu et al. [TCMH11] studied the blinking mechanism in isolation by allowing participants to compare their model generated from real eye-blink data to other commonly used methods for blink animation, such as linear interpolation and ease-in ease-out curves.

Once stimuli have been generated, the order and way in which stimuli are presented must be decided. Most setups show each stimulus under all possible conditions in a random order to each participant (*within-groups* design). An alternative setup was used by Andrist et al. [AMG13], however. In this setup, participants were randomly assigned to one of three gaze aversion conditions in which they performed tasks for each of the proposed hypotheses. This is referred to as the *between-groups* design, where two or more groups of subjects are tested under varied conditions with multiple modified variables simultaneously. In a shared immersive virtual environment, Garau et al. [GSV*03] paired two participants together as part of a role-play, one using a CAVE™-like system and the other using a head-mounted display. Each participant was assigned randomly to one of four conditions to investigate the impact of an avatar's eye

© The Eurographics Association 2014.

gaze behaviour on perceived realism. The setup allowed the authors to test multiple conditions simultaneously and acquire an unbiased opinion from participants on behavioural realism and quality of communication.

Some thought should also be given to where the experiment takes place. While most experiments usually take place in a laboratory due to the available infrastructure, online experiments offer an advantage in that they reach a broader range of subjects and hardware configurations [OSS09]. For researchers located in the U.S., Amazon.com's Mechanical Turk [Ama14] gives researchers the opportunity to access a wide range of potential participants. However, care must be taken to ensure the experiment is set up carefully and that the results are examined critically to avoid selection bias and other confounds [PCI10, BKG11].

6.1.2. Stimuli Presentation and Evaluation

The experiment setup is complete once participant tasks are defined. These assessments use two types of measurements: subjective measurements and objective measurements. Subjective measurements include self-reported assessments of the characters by observers, such as how plausible or likable observers find a character to be. An objective standard, such as how well participants remember information presented by a character, exists for objective measurements.

A user's subjective response to a stimulus can be evaluated by a rating scale, forced choices, a free description task, or an interview. The most common approach to capture personal impressions is a questionnaire employing a rating scale. One such scale is referred to as a *Likert* scale [Lik32], which captures the participant's level of agreement to a statement. The levels used in research literature questionnaires are typically 5- or 7-point scales to ensure an equal number of positive and negative positions and a neutral middle position are incorporated into the scale. McDonnell et al. [MBB12] asked participants to rate different rendered models on a scale from 1 to 7, according to specific definitions (e.g., "extremely abstract-extremely realistic", "unappealing-appealing" and "unfamiliar-familiar"). This enabled them to make a subjective judgement of the participants' experiences on well-defined scales.

An alternative to the Likert scale is the semantic differential method. On the commonly used 7-point scale, the connotative meaning of a concept is measured [OST57]. Adjectives of opposite meaning define extremes at both ends of the scale. A representative example of the semantic differential method is the evaluation of users' impressions on a presented gaze model by Fukayama et al. [FOM*02], which used twenty pairs of polar adjectives adopted from psychological studies, such as "extrovert versus introvert" or "lazy versus diligent". This technique is useful when the experimenter wishes to measure the attitudes of participants to the experimental stimuli. For example, a human motion could be presented and participants asked to judge it in terms of qualities such as happiness, energy level, and so on.

The difficulty with rating scales lies in creating questions with answers that clearly map to all points on the scale. For example, on a 1- to 7- point scale, the participant needs to know what 1 and 7 mean when answering a question, such as "How happy does this character look?". In practice, a number of researchers define new sets of questionnaires and mappings adapted to their requirements. A good resource for concepts to capture the impression of an avatar's personality or naturalness is the field of psychology. The *Ten Item Personality Inventory* (TIPI) [GRS03] was defined to capture the *Big Five personality traits* [Nor63] describing the human personality: openness, conscientiousness, extraversion, agreeableness, and neuroticism. To test the effect on the perception of social dominance in synthesised conversational behaviours, Bee et al. [BPAW10] used TIPI as reference for extraversion and agreeableness. The same study also derived six questions from the *PAD emotional state model* to measure dominance. The psychological PAD model, developed by Mehrabian, represents the fundamental human emotions in the dimensions: *Pleasure*, *Arousal*, and *Dominance* [Meh80]. In connection with the evaluation of emotions, the use of *self-assessment manikins* (SAMs) [GK05] should be mentioned. SAMs is a non-verbal technique that has been used to capture the three PAD dimensions of emotions [BDG*07]. Each dimension is depicted by stylised figures on a 5-point scale. Used for the recognition of emotions in speech, this system has been shown to be simple and straightforward [BDG*07].

In general, rating scales are an effective way of collecting data quickly. However, a disadvantage of the approach is that participants may give socially desirable responses, or they may develop a response set (e.g., giving consistently moderate answers). A forced choice task restricts the responses a user can give to a specific question. Presenting a limited number of cases from which the user chooses avoids the problem of difficulty in interpretation. In the case of evaluating the naturalness of their head-and-eye motion model, Le et al. [LMD12] generated video clips using different methods, including their own, to animate head and eye motion and facial gestures. Two sequential clips were shown to the participant, who was asked to choose the more natural one. Excluding the motion capture animations, user votes were in favour of the speech-driven head and eye motion generator proposed by the authors.

After the experiment has taken place, it may be useful to include a free writing questionnaire or post-experiment interview in order to obtain additional information about participants' opinions. Following an experiment to test the effect of a robot's gaze behaviour on perceived believability and likability, Poel et al. [PHN*09] carried out a semi-structured interview. More direct questions about the user's experience with the robot provided further qualitative infor-

© The Eurographics Association 2014.

mation for comparison between the robot using the gaze behaviour system and the robot not using it.

In contrast to subjective measurements, the objective assessment of a method is possible by using different task designs. Such an objective approach was used by Murray et al. [MRS*07] in an immersive virtual environment where the user had to pick the object from the screen at which he thought the avatar was looking. An objective measurement to verify whether or not the gaze model by Andrist et al. [APMG12a] improves learning took the form of a quiz taken by participants after they watched an agent present a lecture (Figure 13). A virtual agent gave a three minute lecture on a Chinese city visible on a map under conditions employing affiliative gaze, referential gaze, or a combination of both gaze types. The ten questions on the quiz were divided into categories unknown to the user: recalling the knowledge associated with the map, the verbal content of the lecture, and information linked between them. The objective assessment indicated that referential gaze improved learning performance, while the affiliative condition was rated better in the subjective measures.

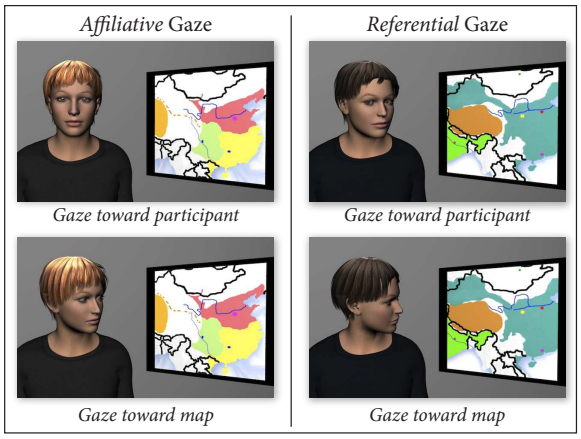

Figure 13: *The characters in this study were delivering 3 minute lectures. Assessment included both subjective questions about rapport with the agent and objective questions about how much of the lecture was remembered. [APMG12a]*

Alternative assessment methods that can be employed to explore the design space in the formative stages of the process have been proposed, although these methods have not yet found common use. An example is a "multivariate" evaluation approach in which a number of design variables, such as gaze frequency at different targets, fixation length, or alignment between the eyes and the head are simultaneously manipulated along a continuous scale, such as x_i ~ U(0,1). Their joint effects on the measured outcomes are then modelled. Huang and Mutlu [HM13] used this approach to assess how different types of arm gestures of a storytelling robot affected how much participants recalled the details of the robot's story, the participants' perceptions of the robot as an effective storyteller, and their ability to retell the story.

Mathematical and computer-based evaluations offer an alternative to human participants. Testing the efficiency and performance improvement of a proposed approach compared to other approaches is one way [GT09]. An example of a computer-based evaluation was carried out by Itti et al. [ID03]. The eye-head movements of their model were traced, demonstrating the basic modes and comparing them to possible human behaviours. Using various video segments as input, the horizontal and vertical eye and head coordinates were plotted on 2D graphs to illustrate the "correct" (human) tracking behaviour.

6.2. Analysis

The final step in evaluating a proposed modelling technique is data analysis. Data reduction, descriptive statistics, or inferential statistics, among other methods, are applicable depending on the task performed. Factor analysis is suited to results from semantic differential rating scales as in case of Fukayama et al. [FOM*02]. For the evaluation of other rating scales, such as the Likert scale, Analysis of Variance (ANOVA) can be used [MB10]. A post-hoc test on the ANOVA result further improves the reliability of the analysis. There are different methods available, such as Scheffé's method for a priori contrast tests [AMG13], the Tukey-Kramer HSD for experiments with unequal sizes of groups, Newman-Keuls tests for comparison of means [ZHRM13], or Bonferroni corrections for multiple comparisons [TCMH11]. Each of these tests is suited better for a given test result and has advantages and disadvantages. While the Tukey and Newman-Keuls tests are more suitable for simple comparison, Scheffé's method or Bonferroni are useful when the experiment consists of subgroups. The choice of the appropriate test therefore depends on the research question, the given dataset, what needs to be compared, and the desired power of the analysis.

As used by Poel et al. [PHN*09], the Wilcoxon signed-rank test [Wil45] is a non-parametric test for comparison of two related samples. The consistency test method by Kendall and Smith [KBS40], adapted from biometrics, was also found to be used as a method to compare the results for different presented methods [MD09]. There is a wide range of analysis techniques to choose from when conducting an experiment. Please see further literature on experimental design for a more in depth discussion on analytic techniques (e.g. [KW04, How09]).

7. Conclusions

In this report, we have presented the accomplishments to date on generating artificial entities that aim to replicate the

© The Eurographics Association 2014.

motion and appearance of the human eye. From our perspective, there are a number of important future directions such as how models can be used in practice between applications, how to create variety and individuality in gaze behaviour, and how best to create more expressivity in cartoon characters.

One critical aspect of future gaze modelling research will be to better connect models to application-level outcomes so that agent developers can use models in practice. Achieving such a transfer will require three tasks: (1) cataloguing the kinds of effects that agents can achieve through proper use of the agents' gaze, (2) providing techniques that synthesise the appropriate gaze behaviours to achieve these effects, and validating that these effects can be achieved through the use of the techniques, and (3) demonstrating potential benefits of the use of designed agent gaze behaviours. While some success has been achieved in illustrating these effects, to date these have been in small, controlled scenarios. For example, Andrist et al. [AMG13] showed improved learning effects for subjects watching brief videos; however, future work should explore larger (lecture or even semester course) scale scenarios and longer term effects such as long term recall.

The majority of eye-gaze methods discussed in this report have attempted to create a generic model of gaze behaviour for all agents. One aspect that has been neglected is the idea of individuality and inter-personal differences in eye gaze. This may be particularly important when simulating groups or crowds of agents in order to increase the variety and individuality of the agents. Previous research has identified the importance of variety in both appearance and motion of agents [MLD*08, MLH*09], but has yet to investigate whether head and eye gaze could affect crowd variety also.

Computational gaze modelling has focused on reproducing the mechanisms seen in real humans and animals. However, animated characters are not subject to the same constraints as the real world and might be able to achieve more effective communication by using different mechanisms. Cartoon characters have long achieved communicative effects using eye movements that are impossible in the real world; for example, spinning eyes, eyes shooting rays at a target, and eyes bulging. Understanding the potential for increased expressiveness and the tradeoffs that arise once paradigms are broken is an interesting avenue for future work.

Although much work has been accomplished in this field, there still remain many challenges ahead before we can replicate the complexity and individuality of human eye behaviour. We have indicated some of the future directions from our perspective, but anticipate a much wider range of future activities in this area.

Acknowledgements

This work was sponsored by Science Foundation Ireland as part of the "Cartoon Motion" project. This research was also supported by the NSF award 1208632 and "ERC advanced grant EXPRESSIVE", the U.S. National Science Foundation awards 1017952 and 1149970, and the U.S. National Institutes of Health award 1R01HD071089-01A1.

Author biographies:

Fitting the cross-disciplinary nature of the subject domain and aims of the survey, the authors are an international team of recognised specialists in eye-gaze from various disciplines. These include neuroscience, psychology, intelligent virtual agents, computer graphics and animation, computer games and serious games, and human-robot interaction. The authors have published in leading conferences and journals such as Computer Graphics Forum, ACM Transactions on Graphics, ACM Transactions on Applied Perception, IEEE System Man and Cybernetics, ACM Transactions on Intelligent Interactive Systems, Experimental Brain Research, etc. The authors list spans from those with experience in the domain of eye-gaze for many years, to newer researchers bringing fresh perspectives and vigour to the domain.

Kerstin Ruhland received her Degree in Computer Science from the Ludwig Maximilian University of Munich in 2009. Since March 2013, she has been working toward a Ph.D degree at Trinity College Dublin, under the supervision of Rachel McDonnell. Her main focus is on realistic facial and eye animation.

Sean Andrist is a Ph.D student working with Dr. Bilge Mutlu in the Human-Computer Interaction lab and Dr. Michael Gleicher in the Visual Computing lab at the University of Wisconsin-Madison. His current research involves the development of communicative characters, including both embodied virtual agents and physical robots, that can use humanlike gaze mechanisms to effectively communicate with people in a variety of contexts, including education, therapy, and collaborative work.

Jeremy B. Badler received his Doctorate in Neuroscience from the University of Louvain in 2011. He is currently a postdoctoral research fellow at the Playable Innovative Technologies Lab at Northeastern University. After beginning his academic life as an engineer, he gradually transitioned into systems-level neuroscience. His previous work has used oculomotor psychophysics, and to a lesser extent single-unit electrophysiology, to try to elucidate the processes behind timing, prediction, anticipation and causality interpretation in the motor system of the brain.

Christopher E. Peters received the Ph.D. in Computer Science from Trinity College Dublin in 2004. He is an Associate Professor of HCI and computer graphics in the School

© The Eurographics Association 2014.

of Computer Science and Communication (CSC) at KTH Royal Institute of Technology, Stockholm. His research interests include the investigation of interrelationships between computational visual attention, affect, memory, theory of mind and gaze control models for real-time animation, perception of, and multimodal interaction with humanoid characters, groups and crowds. He is the co-founder of the AFFINE (Affective Interaction in Natural Environments) events which have taken place at various venues, including ACM ICMI and ACM Multimedia.

Norman I. Badler is the Rachleff Professor of Computer and Information Science at the University of Pennsylvania. He received his Ph.D in Computer Science from the University of Toronto in 1975. His research interests include embodied agent animation and simulation, human-computer interfaces, crowd modelling and control, and computational connections between language and action. He is the founding Director of the SIG Center for Computer Graphics and the Center for Human Modelling and Simulation at Penn.

Michael Gleicher earned his Ph.D. in Computer Science from Carnegie Mellon University, and holds a B.S.E. in Electrical Engineering from Duke University. He is a Professor in the Department of Computer Sciences at the University of Wisconsin, Madison and founder and leader of the Department's Computer Graphics group. His current research includes data visualization, computational videography, character animation, and robotics.

Bilge Mutlu received his Ph.D degree from Carnegie Mellon University Human-Computer Interaction Institute. Currently he is an Assistant Professor of Computer Science, Psychology, and Industrial Engineering at the University of Wisconsin-Madison and the Director of the Wisconsin Human Computer Interaction Laboratory. His main research interests involve Human computer interaction (HCI); human-robot interaction (HRI); computer- supported collaborative work (CSCW); computer-mediated communication (CMC); use of humanlike robots and agents in education, health, collaboration, communication, and research; behaviour and cognitive modelling and brain-computer interfaces.

Rachel McDonnell received her Ph.D. in Computer Science from Trinity College Dublin in 2006 on the topic of crowd animation and perception. She is an Assistant Professor in Creative Technologies at the School of Computer Science and Statistics at Trinity College Dublin. Her research interests include computer graphics, character animation, and perceptually-adaptive graphics. Her particular focus is on emotionally expressive facial animation and perception.

References

[Abe86] ABELE A.: Functions of gaze in social interaction: Communication and monitoring. *Journal of Nonverbal Behavior 10*, 2 (1986), 83–101. 9

[AC76] ARGYLE M., COOK M.: *Gaze and Mutual Gaze*. Cambridge University Press, 1976. 8

[AHS02] ALBRECHT I., HABER J., SEIDEL H.-P.: Automatic generation of non-verbal facial expressions from speech. In *Proc. Computer Graphics International 2002* (2002), pp. 283–293. 10

[Ama14] AMAZON.COM: Amazon mechanical turk. http://www.mturk.com/mturk/. 14

[AMG13] ANDRIST S., MUTLU B., GLEICHER M.: Conversational gaze aversion for virtual agents. In *Intelligent Virtual Agents* (2013), vol. 8108 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 249–262. 9, 13, 15, 16

[And81] ANDERSON J. H.: Ocular torsion in the cat after lesions of the interstitial nucleus of cajal. *Annals of the New York Academy of Sciences 374*, 1 (1981), 865–871. 5

[APMG12a] ANDRIST S., PEJSA T., MUTLU B., GLEICHER M.: Designing effective gaze mechanisms for virtual agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), CHI '12, ACM, pp. 705–714. 7, 13, 15

[APMG12b] ANDRIST S., PEJSA T., MUTLU B., GLEICHER M.: A head-eye coordination model for animating gaze shifts of virtual characters. In *Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction* (2012), Gaze-In '12, pp. 4:1–4:6. 5, 7

[ATGM14] ANDRIST S., TAN X. Z., GLEICHER M., MUTLU B.: Conversational gaze aversion for humanlike robots. In *Proceedings of HRI 2014* (2014). 12

[AWH10] ANDERSON C., WALES A., HORNE J.: PVT lapses differ according to eyes open, closed, or looking away. *Sleep* (2010), 3:197–3:204. 6

[BAP*96] BENTIN S., ALLISON T., PUCE A., PEREZ E., MCCARTHY G.: Electrophysiological studies of face perception in humans. *J. Cognitive Neuroscience 8*, 6 (1996), 551–565. 11

[BAS75] BAHILL A. T., ADLER D., STARK L.: Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology & Visual Science 14*, 6 (1975), 468–9. 4

[BAT09] BEE N., ANDRÉ E., TOBER S.: Breaking the ice in human-agent communication: Eye-gaze based initiation of contact with an embodied conversational agent. In *Intelligent Virtual Agents*, vol. 5773 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2009, pp. 229–242. 10

[BBBL01] BAILENSON J. N., BLASCOVICH J., BEALL A. C., LOOMIS J. M.: Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments 10*, 6 (2001), 583–598. 9

[BBLB02] BLASCOVICH J., BEALL A. C., LOOMIS J. M., BAILENSON J. N.: Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin 29* (2002), 1–15. 9

[BBW96] BURGOON J., BULLER D., WOODALL W.: *Nonverbal Communication: The Unspoken Dialogue*. McGraw-Hill Higher Education, 1996. 6

[BCS75] BAHILL A., CLARK M. R., STARK L.: The main sequence, a tool for studying human eye movements. *Mathematical Biosciences 24*, 3?4 (1975), 191–204. 4

© The Eurographics Association 2014.

[BDG*07] BUSSO C., DENG Z., GRIMM M., NEUMANN U., NARAYANAN S.: Rigid head motion in expressive speech animation: Analysis and synthesis. *IEEE Transactions on Audio, Speech, and Language Processing 15*, 3 (2007), 1075–1086. 14

[BDNN07] BUSSO C., DENG Z., NEUMANN U., NARAYANAN S.: Learning expressive human-like head motion sequences from speech. In *Data-Driven 3D Facial Animation*. Springer London, 2007, pp. 113–131. 9

[Bec89] BECKER W.: The neurobiology of saccadic eye movements. metrics. *Reviews of oculomotor research 3* (1989), 13–67. 4

[Bev13] BEVACQUA E.: A survey of listener behavior and listener models for embodied conversational agents. In *Coverbal Synchrony in Human-Machine Interaction*, Rojc M., Campbell N., (Eds.). Taylor & Francis, 2013, pp. 243–268. 9

[BFJ*05] BENNEWITZ M., FABER F., JOHO D., SCHREIBER M., BEHNKE S.: Integrating vision and speech for conversations with multiple persons. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2005), pp. 2523–2528. 13

[BH10] BOHUS D., HORVITZ E.: Facilitating multiparty dialog with gaze, gesture, and speech. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction* (2010), ICMI-MLMI '10, ACM, pp. 5:1–5:8. 9

[BHAF09] BARD E., HILL R., ARAI M., FOSTER M. E.: Referring and gaze alignment: accessibility is alive and well in situated dialogue. In *Proceedings of CogSci 2009* (2009), Cognitive Science Society, pp. 1246–1251. 9

[BHKS11] BAINBRIDGE W., HART J., KIM E., SCASSELLATI B.: The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics 3*, 1 (2011), 41–52. 12

[BHZS12] BRENNAN S. E., HANNA J. E., ZELINSKY G. J., SAVIETTA K. J.: Eye gaze cues for coordination in collaborative tasks. In *DUET 2012 Workshop: Dual eye tracking in CSCE. 2012 ACM Conference on Computer Supported Cooperative Work* (Seattle, WA, 2012). 9

[BKG11] BUHRMESTER M., KWANG T., GOSLING S. D.: Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science 6*, 1 (2011), 3–5. 14

[BMEL08] BRADLEY M. M., MICCOLI L., ESCRIG M. A., LANG P. J.: The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology 45*, 4 (2008), 602–607. 4

[BPAW10] BEE N., POLLOCK C., ANDRÉ E., WALKER M.: Bossy or wimpy: expressing social dominance by combining gaze and linguistic behaviors. In *Proceedings of the 10th international conference on Intelligent virtual agents* (2010), IVA'10, pp. 265–271. 11, 14

[BPG*06] BRETON G., PELÉ D., GARCIA C., PANAGET F., BRETIER P.: Modeling gaze behavior for a 3d eca in a dialogue situation. In *Gesture in Human-Computer Interaction and Simulation*, Gibet S., Courty N., Kamp J.-F., (Eds.), vol. 3881 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2006, pp. 252–255. 10

[BWA*10] BEE N., WAGNER J., ANDRE E., VOGT T., CHARLES F., PIZZI D., CAVAZZA M.: Gaze behavior during interaction with a virtual character in interactive storytelling. In *AAMAS 2010 Workshop on Interacting with ECAs as Virtual Characters* (2010). 9

[Car78] CARY M. S.: The Role of Gaze in the Initiation of Conversation. *Social Psychology 41*, 3 (1978), 269–271. 10

[Cas00] CASSELL J.: *Embodied Conversational Agents*. MIT Press, 2000. 8

[CCD00] COHEN M. F., COLBURN R. A., DRUCKER S. M.: *The Role of Eye Gaze in Avatar Mediated Conversational Interfaces*. Tech. rep., 2000. 10

[CGV09] CAFARO A., GAITO R., VILHJÁLMSSON H. H.: Animating idle gaze in public places. In *Proceedings of the 9th International Conference on Intelligent Virtual Agents* (2009), IVA '09, Springer-Verlag, pp. 250–256. 8

[Cho92] CHOVIL N.: Discourse-oriented facial displays in conversation. *Research on Language & Social Interaction 25* (1991/1992), 163?194. 10

[CK03] CLARK H. H., KRYCH M. A.: Speaking while monitoring addressees for understanding. *Journal of Memory and Language 50*, 1 (2003), 62–81. 9

[CKB99] CHOPRA-KHULLAR S., BADLER N. I.: Where to look? automating attending behaviors of virtual human characters. In *Proceedings of the third annual conference on Autonomous Agents* (1999), AGENTS '99, ACM, pp. 16–23. 5, 8

[CKEMT10] CIG C., KASAP Z., EGGES A., MAGNENAT-THALMANN N.: Realistic emotional gaze and head behavior generation based on arousal and dominance factors. In *MIG* (2010), vol. 6459 of *Lecture Notes in Computer Science*, Springer, pp. 278–289. 8

[CPB*94] CASSELL J., PELACHAUD C., BADLER N., STEEDMAN M., ACHORN B., BECKET T., DOUVILLE B., PREVOST S., STONE M.: Animated conversation: rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (1994), SIGGRAPH '94, ACM, pp. 413–420. 11

[CT99] CASSELL J., THÓRISSON K. R.: The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence 13*, 4-5 (1999), 519–538. 9

[CT07] CASSELL J., TARTARO A.: Intersubjectivity in human-agent interaction. *Interaction Studies 8*, 3 (2007), 391–410. 12

[CTP98] CASSELL J., TORRES O. E., PREVOST S.: Turn taking vs. discourse structure: How best to model multimodal conversation. In *Machine Conversations* (1998), Kluwer, pp. 143–154. 9

[CV99] CASSELL J., VILHJÁLMSSON H. H.: Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems 2*, 1 (1999), 45–64. 8

[CVB01] CASSELL J., VILHJÁLMSSON H. H., BICKMORE T.: Beat: the behavior expression animation toolkit. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (2001), SIGGRAPH '01, ACM, pp. 477–486. 11

[CVB*12] CAFARO A., VILHJÁLMSSON H., BICKMORE T., HEYLEN D., JÓHANNSDÓTTIR K., VALGARÐSSON G.: First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In *Intelligent Virtual Agents*, vol. 7502 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 67–80. 10

[DdGB10] DELAUNAY F., DE GREEFF J., BELPAEME T.: A study of a retro-projected robotic face and its effectiveness for

© The Eurographics Association 2014.

gaze reading by humans. In *5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2010), pp. 39–44. 12

[DLN05] DENG Z., LEWIS J. P., NEUMANN U.: Automated eye motion using texture synthesis. *IEEE Computer Graphics and Applications 25*, 2 (2005), 24–30. 5, 6, 13

[DLN07] DENG Z., LEWIS J., NEUMANN U.: Realistic eye motion synthesis by texture synthesis. In *Data-Driven 3D Facial Animation*, Deng Z., Neumann U., (Eds.). Springer London, 2007, pp. 98–112. 5, 6

[Dou01] DOUGHTY M. J.: Consideration of Three Types of Spontaneous Eyeblink Activity in Normal Humans: during Reading and Video Display Terminal Use, in Primary Gaze, and while in Conversation. *Optometry and Vision Science 78*, 10 (2001), 712–725. 6

[Dun74] DUNCAN STARKEY J.: On the structure of speaker-auditor interaction during speaking turns. *Language in Society 3*, 2 (1974), 161–180. 10

[EF03] EKMAN P., FRIESEN W.: *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Malor Books, 2003. 1

[EMP*94] EVINGER C., MANNING K. A., PELLEGRINI J. J., BASSO M. A., POWERS A. S., SIBONY P. A.: Not looking while leaping: the linkage of blinking and saccadic gazeshifts. *Experimental brain research 100*, 2 (1994), 337–344. 6

[EMS91] EVINGER C., MANNING K. A., SIBONY P. A.: Eyelid movements. mechanisms and normal data. *Investigative Ophthalmology & Visual Science 32*, 2 (1991), 387–400. 6

[FBT07] FRISCHEN A., BAYLISS A. P., TIPPER S. P.: Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin 133*, 4 (2007), 694. 10

[FGBB09] FRANÇOIS G., GAUTRON P., BRETON G., BOUATOUCH K.: Image-based modeling of the human eye. *IEEE Transactions on Visualization and Computer Graphics 15*, 5 (2009), 815–827. 3

[FN11] FUKUHARA Y., NAKANO Y.: Gaze and conversation dominance in multiparty interaction. In *2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction* (2011). 9

[FOM*02] FUKAYAMA A., OHNO T., MUKAWA N., SAWAKI M., HAGITA N.: Messages embedded in gaze of interface agents – impression management with agent's gaze. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2002), CHI '02, ACM, pp. 41–48. 8, 14, 15

[FR84] FISCHER B., RAMSPERGER E.: Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research 57*, 1 (1984), 191–195. 4

[FTT99] FUNGE J., TU X., TERZOPOULOS D.: Cognitive modeling: Knowledge, reasoning and planning for intelligent characters. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques* (1999), SIGGRAPH '99, ACM Press/Addison-Wesley Publishing Co, pp. 29–38. 8

[Ful92] FULLER J.: Head movement propensity. *Experimental Brain Research 92*, 1 (1992), 152–164. 6

[GB06] GU E., BADLER N. I.: Visual attention and eye gaze during multiparty conversations with distractions. In *Intelligent Virtual Agents*, Gratch J., Young M., Aylett R., Ballin D., Olivier P., (Eds.), vol. 4133 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2006, pp. 193–204. 9

[GC11] GERGLE D., CLARK A. T.: See what I'm saying? using dyadic mobile eye tracking to study collaborative reference. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11* (2011), ACM Press, p. 435. 9

[GD02] GILLIES M. F. P., DODGSON N. A.: Eye movements and attention for behavioural animation. *Journal of Visualization and Computer Animation 13*, 5 (2002), 287–300. 8

[GK05] GRIMM M., KROSCHEL K.: Evaluation of natural emotions using self assessment manikins. In *IEEE Workshop on Automatic Speech Recognition and Understanding* (2005), pp. 381–385. 14

[GLBB07] GU E., LEE S., BADLER J. B., BADLER N. I.: Eye movements, saccades, and multiparty conversations. In *Data-Driven 3D Facial Animation*. Springer London, 2007, pp. 79–97. 5, 6, 9

[Gof63] GOFFMAN E.: *Behaviour in public places: notes on the social order of gatherings*. The Free Press, New York, 1963. 10

[GP09] GATICA-PEREZ D.: Modeling interest in face-to-face conversations from multimodal nonverbal behaviour. In *Multi-Modal Signal Processing: Methods and Techniques to Build Multimodal Interactive Systems*, Thiran J., Bourlard H., Marqués F., (Eds.), EURASIP and Academic Press series in signal and image processing. Elsevier Science, 2009, pp. 309–326. 9

[Gre86] GREENE P. R.: Gaussian and poisson blink statistics: A preliminary study. *IEEE Transactions on Biomedical Engineering BME-33*, 3 (1986), 359–361. 6

[GRS03] GOSLING S. D., RENTFROW P. J., SWANN W. B.: A very brief measure of the big-five personality domains. *Journal of Research in Personality 37*, 6 (2003), 504–528. 14

[GSC91] GUITTON D., SIMARD R., CODÈRE F.: Upper eyelid movements measured with a search coil during blinks and vertical saccades. *Investigative Ophthalmology & Visual Science 32*, 13 (1991), 3298–3305. 6

[GSV*03] GARAU M., SLATER M., VINAYAGAMOORTHY V., BROGNI A., STEED A., SASSE M. A.: 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* (2003), CHI '03, ACM, pp. 529–536. 13

[GT09] GRILLON H., THALMANN D.: Simulating gaze attention behaviors for crowds. *Computer Animation and Virtual Worlds 20*, 2 (2009), 111–119. 15

[GV87] GUITTON D., VOLLE M.: Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range. *Journal of Neurophysiology 58*, 3 (1987), 427–459. 6, 7

[GWG*07] GRATCH J., WANG N., GERTEN J., FAST E., DUFFY R.: Creating rapport with virtual agents. In *7th International Conference on Intelligent Virtual Agents. 2007* (2007), Springer, pp. 125–138. 9

[HB07] HANNA J. E., BRENNAN S. E.: Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language 57*, 4 (2007), 596–615. 9

[Hey06] HEYLEN D.: Head gestures, gaze and the principles of conversational structure. *International journal of Humanoid Robotics 3*, 3 (2006), 241–267. 9

[HJO*10] HODGINS J., JÖRG S., O'SULLIVAN C., PARK S. I., MAHLER M.: The saliency of anomalies in animated human characters. *ACM Transactions on Applied Perception 7*, 4 (2010), 22:1–22:14. 5

© The Eurographics Association 2014.

[HM13] HUANG C.-M., MUTLU B.: Modeling and evaluating narrative gestures for humanlike robots. In *Proceedings of Robotics: Science and Systems* (2013). 15

[HNP07] HEYLEN D., NIJHOLT A., POEL M.: Generating nonverbal signals for a sensitive artificial listener. In *Verbal and Nonverbal Communication Behaviours*, vol. 4775 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007, pp. 264–274. 9

[HO12] HJALMARSSON A., OERTEL C.: Gaze direction as a back-channel inviting cue in dialogue. In *IVA 2012 workshop on Realtime Conversational Virtual Agents* (2012). 9

[How09] HOWELL D.: *Statistical Methods for Psychology*. Cengage Learning, 2009. 15

[ID03] ITTI L., DHAVALE N.: Realistic avatar eye and head animation using a neurobiological model of visual attention. In *Proceedings of SPIE* (2003), SPIE Press, pp. 64–78. 3, 8, 15

[IDP06] ITTI L., DHAVALE N., PIGHIN F. H.: Photorealistic attention-based gaze animation. In *IEEE International Conference on Multimedia and Expo* (2006), IEEE Press, pp. 521–524. 8

[IMFN06] ISHII R., MIYAJIMA T., FUJITA K., NAKANO Y.: Avatar's gaze control to facilitate conversational turn-taking in virtual-space multi-user voice chat system. In *Intelligent Virtual Agents*, Gratch J., Young M., Aylett R., Ballin D., Olivier P., (Eds.), vol. 4133 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2006, pp. 458–458. 10

[IONN13] ISHII R., OOKO R., NAKANO Y., NISHIDA T.: Effectiveness of gaze-based engagement estimation in conversational agents. In *Eye Gaze in Intelligent User Interfaces*, Nakano Y. I., Conati C., Bader T., (Eds.). Springer London, 2013, pp. 85–110. 10

[Itt00] ITTI L.: *Models of Bottom-Up and Top-Down Visual Attention*. PhD thesis, California Institute of Technology, 2000. 8

[Iza91] IZARD C.: *The Psychology of Emotions*. Emotions, Personality, and Psychotherapy. Springer, 1991. 8

[JHM*07] JAN D., HERRERA D., MARTINOVSKI B., NOVICK D. G., TRAUM D. R.: A computational model of culture-specific conversational behavior. In *IVA* (2007), vol. 4722 of *Lecture Notes in Computer Science*, Springer, pp. 45–56. 9

[JTC*07] JOHNS M., TUCKER A., CHAPMAN R., CROWLEY K., MICHAEL N.: Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. *Somnologie - Schlafforschung und Schlafmedizin 11*, 4 (2007), 234–242. 6

[KB04] KIDD C., BREAZEAL C.: Effect of a robot on user perceptions. In *Proceedings of IROS 2004* (2004), vol. 4, pp. 3559–3564. 12

[KBS40] KENDALL M. G., BABINGTON SMITH B.: On the method of paired comparisons. *Biometrika 31*, 3/4 (1940), 324–345. 15

[Ken67] KENDON A.: Some functions of gaze-direction in social interaction. *Acta psychologica 26*, 1 (1967), 22–63. 9

[Ken90] KENDON A.: *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Studies in Interactional Sociolinguistics. Cambridge University Press, 1990. 8

[KHJK13] KOMOGORTSEV O., HOLLAND C., JAYARATHNA S., KARPOV A.: 2d linear oculomotor plant mathematical model: Verification and biometric applications. *ACM Transactions on Applied Perception 10*, 4 (2013), 27:1–27:18. 5

[KK13] KULMS P., KOPP S.: Using virtual agents to guide attention in multi-task scenarios. In *Intelligent Virtual Agents*, vol. 8108 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 295–302. 10

[KKM*06] KOPP S., KRENN B., MARSELLA S., MARSHALL A. N., PELACHAUD C., PIRKER H., THÓRISSON K. R., VILHJÁLMSSON H.: Towards a common framework for multimodal generation: The behavior markup language. In *Proceedings of the 6th International Conference on Intelligent Virtual Agents* (2006), IVA'06, Springer-Verlag, pp. 205–217. 11

[Kle86] KLEINKE C. L.: Gaze and eye contact: a research review. *Psychological Bulletin 100*, 1 (1986), 78–100. 9

[KOS11] KOKKINARA E., OYEKOYA O., STEED A.: Modelling selective visual attention for autonomous virtual characters. *Computer Animation and Virtual Worlds 22*, 4 (2011), 361–369. 8

[KPPP11] KRENN B., PELACHAUD C., PIRKER H., PETERS C.: Embodied conversational characters: Representation formats for multimodal communicative behaviours. In *Emotion-Oriented Systems*, Cowie R., Pelachaud C., Petta P., (Eds.), Cognitive Technologies. Springer Berlin Heidelberg, 2011, pp. 389–415. 11

[KW04] KEPPEL G., WICKENS T.: *Design and Analysis: A Researcher's Handbook*. Prentice Hall, 2004. 15

[KYH*98] KIKUCHI H., YOKOYAMA M., HOASHI K., HIDAKI Y., KOBAYASHI T., SHIRAI K.: Controlling gaze of humanoid in communication with human. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (1998), vol. 1, pp. 255–260. 13

[LB06] LAM M. W., BARANOSKI G. V.: A predictive light transport model for the human iris. *Computer Graphics Forum 25*, 3 (2006), 359–368. 3

[LBB02] LEE S. P., BADLER J. B., BADLER N. I.: Eyes alive. *ACM Transactions on Graphics 21*, 3 (2002), 637–644. 5, 6

[LBS*03] LEFOHN A., BUDGE B., SHIRLEY P., CARUSO R., REINHARD E.: An ocularist's approach to human iris synthesis. *IEEE Computer Graphics and Applications 23*, 6 (2003), 70–75. 3

[Lik32] LIKERT R.: A technique for the measurement of attitudes. *Archives of Psychology 22*, 140 (1932), 1–55. 14

[LM07] LANCE B., MARSELLA S.: Emotionally expressive head and body movement during gaze shifts. In *IVA* (2007), vol. 4722 of *Lecture Notes in Computer Science*, Springer, pp. 72–85. 8

[LM10a] LANCE B., MARSELLA S.: The expressive gaze model: Using gaze to express emotion. *IEEE Computer Graphics and Applications 30*, 4 (2010), 62–73. 8

[LM10b] LANCE B., MARSELLA S.: Glances, glares, and glowering: how should a virtual human express emotion through gaze? *Autonomous Agents and Multi-Agent Systems 20*, 1 (2010), 50–69. 8

[LM12] LI Z., MAO X.: Emotional eye movement generation based on geneva emotion wheel for virtual agents. *J. Vis. Lang. Comput 23*, 5 (2012), 299–310. 9

[LMD12] LE B., MA X., DENG Z.: Live speech driven head-and-eye motion generators. *IEEE Transactions on Visualization and Computer Graphics 18*, 11 (2012), 1902–1914. 5, 6, 7, 10, 14

[LMK04] LANCE B., MARSELLA S., KOIZUMI D.: Towards expressive gaze manner in embodied virtual agents. In *Autonomous Agents and Multi-Agent Systems Workshop on Empathic Agents* (2004). 8

[LMT*07] LEE J., MARSELLA S., TRAUM D., GRATCH J., LANCE B.: The rickel gaze model: A window on the mind of

© The Eurographics Association 2014.

a virtual human. In *Proceedings of the 7th international conference on Intelligent Virtual Agents* (2007), IVA '07, Springer-Verlag, pp. 296–303. 8

[LR86] LAURUTIS V. P., ROBINSON D. A.: The vestibulo-ocular reflex during human saccadic eye movements. *The Journal of Physiology 373*, 1 (1986), 209–233. 6, 7

[LvW12] LOHSE M., VAN WELBERGEN H.: Designing appropriate feedback for virtual agents and robots. In *Position paper at RO-MAN 2012 Workshop "Robot Feedback in Human-Robot Interaction: How to Make a Robot "Readable" for a Human Interaction Partner"* (2012). 12

[LZ99] LEIGH R. J., ZEE D. S.: *The neurology of eye movements*, 3 ed. No. 55 in Contemporary neurology series. Oxford University Press, 1999. 4, 5, 6

[MB10] MCDONNELL R., BREIDT M.: Face reality: Investigating the uncanny valley for virtual faces. In *ACM SIGGRAPH ASIA 2010 Sketches* (2010), SA '10, ACM, pp. 41:1–41:2. 15

[MB12] MARIOORYAD S., BUSSO C.: Generating human-like behaviors using joint, speech-driven models for conversational agents. *IEEE Transactions on Audio, Speech, and Language Processing 20*, 8 (2012), 2329–2340. 10

[MBB12] MCDONNELL R., BREIDT M., BÜLTHOFF H. H.: Render me real?: investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics 31*, 4 (2012), 91:1–91:11. 5, 14

[MD09] MA X., DENG Z.: Natural eye motion synthesis by modeling gaze-head coupling. In *Proceedings of the 2009 IEEE Virtual Reality Conference* (2009), VR '09, IEEE Computer Society, pp. 143–150. 15

[MDR*11] MARTIN J.-C., DEVILLERS L., RAOUZAIOU A., CARIDAKIS G., RUTTKAY Z., PELACHAUD C., MANCINI M., NIEWIADOMSKI R., PIRKER H., KRENN B., POGGI I., CALDOGNETTO E., CAVICCHIO F., MEROLA G., ROJAS A., VEXO F., THALMANN D., EGGES A., MAGNENAT-THALMANN N.: Coordinating the generation of signs in multiple modalities in an affective agent. In *Emotion-Oriented Systems*, Cowie R., Pelachaud C., Petta P., (Eds.), Cognitive Technologies. Springer Berlin Heidelberg, 2011, pp. 349–367. 11

[MEB12] MOUBAYED S. A., EDLUND J., BESKOW J.: Taming mona lisa: Communicating gaze faithfully in 2d and 3d facial projections. *ACM Transactions on Interactive Intelligent Systems 1*, 2 (2012), 11:1–11:25. 12

[Meh80] MEHRABIAN A.: *Basic Dimensions for a General Psychological Theory*. OGH Publishers, 1980. 14

[MGR04] MARSELLA S., GRATCH J., RICKEL J.: Expressive behaviors for virtual worlds. In *Life-like characters*, Cognitive Technologies. Springer, 2004, pp. 317–360. 8

[MH07] MASUKO S., HOSHINO J.: Head-eye animation corresponding to a conversation for cg characters. *Computer Graphics Forum 26*, 3 (2007), 303–312. 5, 6, 7

[MHKS07] MITAKE H., HASEGAWA S., KOIKE Y., SATO M.: Reactive virtual human with bottom-up and top-down visual attention for gaze generation in realtime interactions. In *Virtual Reality Conference, 2007. VR '07. IEEE* (2007), pp. 211–214. 8

[MLD*08] MCDONNELL R., LARKIN M., DOBBYN S., COLLINS S., O'SULLIVAN C.: Clone attack! perception of crowd variety. *ACM Transactions on Graphics 27*, 3 (2008), 26:1–26:8. 16

[MLH*09] MCDONNELL R., LARKIN M., HERNÁNDEZ B., RUDOMIN I., O'SULLIVAN C.: Eye-catching crowds: Saliency based selective variation. *ACM Transactions on Graphics 28*, 3 (2009), 55:1–55:10. 16

[MM11] MUMM J., MUTLU B.: Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th international conference on Human-Robot Interaction* (2011), HRI '11, ACM, pp. 331–338. 13

[Mor70] MORI M.: The uncanny valley. *Energy 7* (1970), 33–35. 1

[MRS*07] MURRAY N., ROBERTS D., STEED A., SHARKEY P., DICKERSON P., RAE J.: An assessment of eye-gaze potential within immersive virtual environments. *ACM Transactions on Multimedia Computing, Communications, and Applications 3*, 4 (2007), 8:1–8:17. 15

[MSH*06] MOJZISCH A., SCHILBACH L., HELMERT J. R., PANNASCH S., VELICHKOVSKY B. M., VOGELEY K.: The effects of self-involvement on attention, arousal, and facial expression during social interaction with virtual others: A psychophysiological study. *Social Neuroscience 1*, 3-4 (2006), 184–195. 10

[MSK*09] MUTLU B., SHIWA T., KANDA T., ISHIGURO H., HAGITA N.: Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE international conference on Human-Robot Interaction* (2009), HRI '09, ACM, pp. 61–68. 13

[MSSSB10] MARTINEZ S., SLOAN R. J. S., SZYMKOWIAK A., SCOTT-BROWN K. C.: Using virtual agents to cue observer attention. *In: CONTENT 2010: The Second International Conference on Creative Content Technologies* (2010), 7–12. 10

[MVC*12] MCKEOWN G., VALSTAR M., COWIE R., PANTIC M., SCHRODER M.: The semaine database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE Transactions on Affective Computing 3*, 1 (2012), 5–17. 9

[MXL*13] MARSELLA S., XU Y., LHOMMET M., FENG A., SCHERER S., SHAPIRO A.: Virtual character performance from speech. In *Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2013), SCA '13, ACM, pp. 25–35. 10

[NBF*13] NORMOYLE A., BADLER J. B., FAN T., BADLER N. I., CASSOL V. J., MUSSE S. R.: Evaluating perceived trust from procedurally animated gaze. In *Proceedings of Motion on Games* (2013), MIG '13, ACM, pp. 119:141–119:148. 5, 6, 7

[NHP13] NIEWIADOMSKI R., HYNIEWSKA S., PELACHAUD C.: *Computational Models of Expressive Behaviors for a Virtual Agent*. Oxford series on cognitive models and architecture. OUP USA, 2013. 11

[NI10] NAKANO Y. I., ISHII R.: Estimating user's engagement from eye-gaze behaviors in human-agent conversations. In *Proceedings of the 15th International Conference on Intelligent User Interfaces* (2010), IUI '10, ACM, pp. 139–148. 10

[NK10] NAKANO T., KITAZAWA S.: Eyeblink entrainment at breakpoints of speech. *Experimental Brain Research 205*, 4 (2010), 577–581. 6

[NKM*13] NAKANO T., KATO M., MORITO Y., ITOI S., KITAZAWA S.: Blink-related momentary activation of the default mode network while viewing videos. *Proceedings of the National Academy of Sciences 110*, 2 (2013), 702–706. 6

[Nor63] NORMAN W. T.: Toward an adequate taxonomy of personality attributes: replicated factors structure in peer nomination personality ratings. *Journal of abnormal and social psychology 66* (1963), 574–583. 14

[ODK*12] OBAID M., DAMIAN I., KISTLER F., ENDRASS B., WAGNER J., ANDRÉ E.: Cultural behaviors of virtual agents in an augmented reality environment. In *IVA* (2012), vol. 7502 of *Lecture Notes in Computer Science*, Springer, pp. 412–418. 9

© The Eurographics Association 2014.

[OO80] OTTESON J., OTTESON C.: Effect of teacher's gaze on children's story recall. *Perceptual and Motor Skills* (1980). 9

[Opp86] OPPENHEIMER P. E.: Real time design and animation of fractal plants and trees. *ACM SIGGRAPH Computer Graphics 20*, 4 (1986), 55–64. 3

[OSP11] OYEKOYA O., STEED A., PAN X.: Short paper: Exploring the object relevance of a gaze animation model. In *Proceedings of the 17th Eurographics Conference on Virtual Environments & Third Joint Virtual Reality* (2011), EGVE - JVRC'11, Eurographics Association, pp. 111–114. 8

[OSS09] OYEKOYA O., STEPTOE W., STEED A.: A saliency-based method of simulating visual attention in virtual scenes. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology* (2009), VRST '09, ACM, pp. 199–206. 8, 14

[OST57] OSGOOD C. E., SUCI G. J., TANNENBAUM P. H.: *The Measurement of Meaning*. University of Illinois Press, 1957. 14

[PAK09] PETERS C., ASTERIADIS S., KARPOUZIS K.: Investigating shared attention with a virtual agent using a gaze-based interface. *Journal on Multimodal User Interfaces 3*, 1-2 (2009). 10

[PB03] PELACHAUD C., BILVI M.: Modelling gaze behavior for conversational agents. In *Intelligent Virtual Agents (IVA)* (2003), vol. 2792 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 93–100. 9

[PBER07] PICOT A., BAILLY G., ELISEI F., RAIDT S.: Scrutinizing natural scenes: Controlling the gaze of an embodied conversational agent. In *Intelligent Virtual Agents*, vol. 4722 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007, pp. 272–282. 8

[PCI10] PAOLACCI G., CHANDLER J., IPEIROTIS P. G.: Running experiments on amazon mechanical turk. *Judgment and Decision Making 5*, 5 (2010), 411–419. 14

[PCR*11] PETERS C., CASTELLANO G., REHM M., ANDRÉ E., RAOUZAIOU A., RAPANTZIKOS K., KARPOUZIS K., VOLPE G., CAMURRI A., VASALOU A.: Fundamentals of agent perception and attention modelling. In *Emotion-Oriented Systems*, Cognitive Technologies. Springer Berlin Heidelberg, 2011, pp. 293–319. 8

[Pet05] PETERS C.: Direction of attention perception for conversation initiation invirtual environments. In *Intelligent Virtual Agents*, vol. 3661 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2005, pp. 215–228. 10

[Pet06] PETERS C.: Evaluating perception of interaction initiation in virtual environments using humanoid agents. In *Proceedings of the 2006 conference on ECAI 2006: 17th European Conference on Artificial Intelligence August 29* (2006), IOS Press, pp. 46–50. 10

[Pet10] PETERS C.: Animating gaze shifts for virtual characters based on head movement propensity. In *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2010 Second International Conference on* (2010), pp. 11–18. 6, 7

[PHN*09] POEL M., HEYLEN D., NIJHOLT A., MEULEMANS M., VAN BREEMEN A.: Gaze behaviour, believability, likability and the icat. *AI & Society 24*, 1 (2009), 61–73. 14, 15

[PKFT07] POWERS A., KIESLER S., FUSSELL S., TORREY C.: Comparing a computer agent with a humanoid robot. In *Proceedings of HRI 2007* (2007), pp. 145–152. 12

[PMG13] PEJSA T., MUTLU B., GLEICHER M.: Stylized and performative gaze for character animation. *Computer Graphics Forum 32*, 2 (2013), 143–152. 11, 12

[POB09] PAMPLONA V. F., OLIVEIRA M. M., BARANOSKI G. V. G.: Photorealistic models for pupil light reflex and iridal pattern deformation. *ACM Transactions on Graphics 28*, 4 (2009), 106:1–106:12. 3

[POS03] PETERS C., O' SULLIVAN C.: Bottom-up visual attention for virtual human animation. In *Proceedings of the 16th International Conference on Computer Animation and Social Agents (CASA 2003)* (2003), CASA '03, IEEE Computer Society, pp. 111–117. 8

[PPB*05] PETERS C., PELACHAUD C., BEVACQUA E., MANCINI M., POGGI I.: A model of attention and interest using gaze behavior. In *Intelligent Virtual Agents*, vol. 3661 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2005, pp. 229–240. 9

[PPDR00] POGGI I., PELACHAUD C., DE ROSIS F.: Eye communication in a conversational 3d synthetic agent. *AI Commun. 13*, 3 (2000), 169–181. 10

[PQ10] PETERS C., QURESHI A.: Graphics for serious games: A head movement propensity model foranimating gaze shifts and blinks of virtual characters. *Computers and Graphics 34*, 6 (2010), 677–687. 13

[QBM07] QUEIROZ R., BARROS L., MUSSE S.: Automatic generation of expressive gaze in virtual animated characters: From artists craft to a behavioral animation model. In *Intelligent Virtual Agents*, vol. 4722 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007, pp. 401–402. 8

[QBM08] QUEIROZ R. B., BARROS L. M., MUSSE S. R.: Providing expressive gaze to virtual animated characters in interactive applications. *Computers in Entertainment 6*, 3 (2008), 41:1–41:23. 5, 8

[Rem11] REMINGTON L.: *Clinical Anatomy and Physiology of the Visual System*. Elsevier/Butterworth-Heinemann, 2011. 2

[SB12] SCHULMAN D., BICKMORE T.: Changes in verbal and nonverbal conversational behavior in long-term interaction. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction* (2012), ICMI '12, ACM, pp. 11–18. 9

[SBMH94] SAGAR M. A., BULLIVANT D., MALLINSON G. D., HUNTER P. J.: A virtual environment and model of the eye for surgical simulation. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (1994), SIGGRAPH '94, ACM, pp. 205–212. 3

[SC09] STAUDTE M., CROCKER M.: The effect of robot gaze on processing robot utterances. In *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (2009). 13

[Sha11] SHAPIRO A.: Building a character animation system. In *Proceedings of the 4th International Conference on Motion in Games* (2011), MIG'11, Springer-Verlag, pp. 98–109. 5

[Sim94] SIMON B.-C.: How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition 13*, 5 (1994), 513–552. 2

[SM11] SRINIVASAN V., MURPHY R.: A survey of social gaze. In *Proceedings of HRI 2011* (2011), pp. 253–254. 12

[SNJ*07] SKOTTE J., NØJGAARD J., JØRGENSEN L., CHRISTENSEN K., SJØGAARD G.: Eye blink frequency during different computer tasks quantified by electrooculography. *European Journal of Applied Physiology 99*, 2 (2007), 113–119. 6

[SOS10] STEPTOE W., OYEKOYA O., STEED A.: Eyelid kinematics for virtual characters. *Computer Animation and Virtual Worlds 21*, 3-4 (2010), 161–171. 6, 13

© The Eurographics Association 2014.

[SS08] STEPTOE W., STEED A.: High-fidelity avatar eye-representation. In *Virtual Reality Conference, 2008. VR '08. IEEE* (2008), pp. 111–114. 13

[Sta99] STAHL J. S.: Amplitude of human head movements associated with horizontal saccades. *Experimental Brain Research 126*, 1 (1999), 41–54. 6

[SWG84] STERN J. A., WALRATH L. C., GOLDSTEIN R.: The endogenous eyeblink. *Psychophysiology 21*, 1 (1984), 22–33. 6

[TCMH11] TRUTOIU L. C., CARTER E. J., MATTHEWS I., HODGINS J. K.: Modeling and animating eye blinks. *ACM Transactions on Applied Perception 8* (2011), 17:1–17:17. 6, 13, 15

[TL86] TYCHSEN L., LISBERGER S. G.: Visual motion processing for the initiation of smooth-pursuit eye movements in humans. *J Neurophysiol. 56*, 4 (1986), 953–968. 5

[TLM09] THIEBAUX M., LANCE B., MARSELLA S.: Real-time expressive gaze animation for virtual humans. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems* (2009), AAMAS '09, pp. 321–328. 8

[TS04] TOMBS S., SILVERMAN I.: Pupillometry: A sexual selection approach. *Evolution and Human Behavior 25*, 4 (2004), 221–228. 4

[TT94] TU X., TERZOPOULOS D.: Artificial fishes: Physics, locomotion, perception, behavior. In *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques* (1994), SIGGRAPH '94, ACM, pp. 43–50. 8

[VBP11] VALA M., BLANCO G., PAIVA A.: Providing gender to embodied conversational agents. In *IVA* (2011), vol. 6895 of *Lecture Notes in Computer Science*, Springer, pp. 148–154. 9

[VBR*03] VANDERWERF F., BRASSINGA P., REITS D., ARAMIDEH M., ONGERBOER DE VISSER B.: Eyelid movements: behavioral studies of blinking in humans under different stimulus conditions. *J Neurophysiol 89*, 5 (2003), 2784–2796. 6

[VC98] VILHJÁLMSSON H. H., CASSELL J.: Bodychat: Autonomous communicative behaviors in avatars. In *Proceedings of the Second International Conference on Autonomous Agents* (1998), AGENTS '98, ACM, pp. 269–276. 11

[VCC*07] VILHJÁLMSSON H., CANTELMO N., CASSELL J., E. CHAFAI N., KIPP M., KOPP S., MANCINI M., MARSELLA S., MARSHALL A. N., PELACHAUD C., RUTTKAY Z., THÓRISSON K. R., WELBERGEN H., WERF R. J.: The behavior markup language: Recent developments and challenges. In *Proceedings of the 7th International Conference on Intelligent Virtual Agents* (2007), IVA '07, Springer-Verlag, pp. 99–111. 11

[VGSS04] VINAYAGAMOORTHY V., GARAU M., STEED A., SLATER M.: An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience. *Computer Graphics Forum 23*, 1 (2004), 1–12. 5, 7, 10

[Vil04] VILHJÁLMSSON H.: Animating conversation in online games. In *Entertainment Computing ? ICEC 2004*, Rauterberg M., (Ed.), vol. 3166 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2004, pp. 139–150. 11

[WG10] WANG N., GRATCH J.: Don't just stare at me! In *Proceedings of the 28th international conference on Human factors in computing systems* (2010), ACM, pp. 1241–1250. 9

[Wil45] WILCOXON F.: Individual comparisons by ranking methods. *Biometrics Bulletin 1*, 6 (1945), 80–83. 15

[WLO10] WEISSENFELD A., LIU K., OSTERMANN J.: Video-realistic image-based eye animation via statistically driven state machines. *The Visual Computer: International Journal of Computer Graphics 26*, 9 (2010), 1201–1216. 3, 5, 6

[WSG05] WECKER L., SAMAVATI F., GAVRILOVA M.: Iris synthesis: A reverse subdivision application. In *Proceedings of the 3rd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia* (2005), GRAPHITE '05, ACM, pp. 121–125. 3

[XLW13] XU Q., LI L., WANG G.: Designing engagement-aware agents for multiparty conversations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), CHI '13, ACM, pp. 2233–2242. 13

[YLNP12] YEO S. H., LESMANA M., NEOG D. R., PAI D. K.: Eyecatch: simulating visuomotor coordination for object interception. *ACM Transactions on Graphics 31*, 4 (2012), 42:1–42:10. 5, 7

[YNG70] YNGVE V. H.: On getting a word in edgewise. *Chicago Linguistics Society, 6th Meeting* (1970), 567–578. 9

[YSI*06] YOSHIKAWA Y., SHINOZAWA K., ISHIGURO H., HAGITA N., MIYAMOTO T.: Responsive robot gaze to interaction partner. In *Proceedings of robotics: science and systems* (2006). 13

[ZFP11] ZORIC G., FORCHHEIMER R., PANDZIC I.: On creating multimodal virtual humans−real time speech driven facial gesturing. *Multimedia Tools and Applications 54*, 1 (2011), 165–179. 10

[ZHRM13] ZIBREK K., HOYET L., RUHLAND K., MCDONNELL R.: Evaluating the effect of emotion on gender recognition in virtual humans. In *Proceedings of the ACM Symposium on Applied Perception* (2013), SAP '13, ACM, pp. 45–49. 15

[ZS06] ZUO J., SCHMID N. A.: A model based, anatomy based method for synthesizing iris images. In *Proceedings of the 2006 International Conference on Advances in Biometrics* (2006), ICB'06, Springer-Verlag, pp. 428–435. 3

© The Eurographics Association 2014.