Evaluating a Mobile Spontaneous Eye Blink Tracker for use in Tele-presentation HRI as a Low Bandwidth Social Communicative Cue

Chris Bevan and Danaë Stanton Fraser

Abstract—Research suggests that the rate at which humans spontaneously blink their eyes over time is strongly related to their underlying cognitive state. The ability to present the real time blinking behaviour of a human teleoperator via a robot proxy therefore potentially offers observers a low bandwidth - yet salient - cue as to the cognitive state of the teleoperator.

In a controlled study, we demonstrate and evaluate a wireless eye blink detector embedded in a Google Glass wearable computer, transmitting captured blink events in real time for display on a NAO robot. From our evaluation, we present accuracy rates from 28 participants under a range of environmental conditions, describing issues and phenomena encountered.

From a total of 3722 blink events, our prototype blink capture system achieved an overall accuracy of 80.5% across three activity conditions of rest, reading and interview. Additionally, we observed that people blink approximately 33% more frequently when they are listening compared to when they are speaking.

Results are discussed in terms of the requirements of a spontaneous blink detector suitable for capturing real time blinking behaviours within real world conditions.

I. INTRODUCTION

As [28] observes, there is no social interaction without social signals. As robots advance to become socially independent, there is a renewed need to consider how social communicative cues are projected and understood from the perspective of both human and robot. Advances in the relatively recent research domain of social signal processing (e.g. [27]) provides a strong example of how research is responding to this call.

Humans monitor facial cues extensively to judge the emotional / cognitive state of other humans. Given their importance in aiding social interaction, designers of social robots have long recognised the need to incorporate into their designs modulations of human facial cues for expression and have done so in various ways. Blinking lights for example have been used to express emotion and to communicate internal system state for simpler robots (e.g. [14], [9]), while more complex designs (e.g. Hanson Robotics’ Flubber [10] and the Geminoid series of androids [21]) have sought to mimic facial movements in minute detail through replication of the underlying musculature of the human face.

In tele-present HRI, a robot platform may not necessarily possess the ability to display the face of their human operator. In such scenarios, new methods of capturing and projecting the emotional and cognitive state of the operator in real time are needed. Recent research has suggested that spontaneous eye blinking rates (sEBR) have social communicative value [29]. In this paper therefore, we suggest that sEBR’s captured from a robot teleoperator and presented to an observer via the robot in real time may offer a low cost and low bandwidth means of projecting the teleoperator’s internal emotional / cognitive state.

A. Spontaneous Eye Blinking in Humans

An eye blink is a semi-autonomous rapid closure of both eyelids. The duration of a single eye blink typically ranges between 100 - 400 milliseconds, with a highly variable inter-eye-blink interval (IEBI) of between 2-10 seconds [22]. At its most basic level, the eye blink behaviour serves the simple physiological purpose of maintaining lubrication and hydration of the eyeball surface by transferring tear fluid from the inner eyelid. Detailed zoological study has established that nearly all vertebrate animals have closable eyelids that serve this purpose [2]. Eye blinks also occur reflexively as a response to protect the eye from acute invasion by foreign bodies and irritants, and a subtype of reflexive blink described as a startle blink (e.g. [7]) frequently occurs in response to sudden exposure to loud sounds.

The rate at which humans spontaneously blink their eyes however occurs much more frequently - and varies in rate far more widely - than is required to sustain the physiological and defensive needs of the eyeball alone (e.g. [8]). Studies of sEBR report that people blink their eyes within a general range of 8 and 21 blinks per minute [5], with an average of around 17 blinks per minute [1]. Although very young infants are known to blink their eyes considerably less frequently than adults, this phenomenon disappears quickly. Post adolescence, age does not appear to affect blinking rates at all [1]. Likewise, differences in blink rate across biological gender - though inconsistent across studies - are now generally considered to be minimal, with differences observed only when reading [1].

B. The Relationship Between Spontaneous Eye Blinking Rates and Cognition

In a very early study of human blinking behaviour, [20] noted that eye blinking rates appeared to be closely related to the "mental tension" of the participant at the time. A range of studies across disciplines has since established that the rate at which people blink their eyes is closely linked to their underlying cognitive state (e.g. [11], [12]), with a strong relationship between blinking rates, attention and activity. sEBR’s, it has been shown, change drastically with changes
in internal state including arousal, emotion, visuo-attentional demand, fatigue and general cognitive load (e.g. [11]). Relative to rest, sEBR’s have been consistently demonstrated to rise considerably during interpersonal interactions such as conversation (up to 26 blinks / min) and to decrease considerably when reading (as low as 3-4 blinks / min) [1]. A useful heuristic is that as the visual attention demanded by a given task increases, fewer spontaneous blinks will occur. However, as physical and / or cognitive fatigue increases, blinking rates will likely increase.

As a semi autonomous behaviour, spontaneous eye blinking occurs largely unconsciously. The rate at which blinking occurs is determined centrally by the globus pallidus or ‘blinking centre’ of the brain, a deep brain structure in the basal ganglia region widely thought to be involved in the regulation of movement. During fluid social interpersonal interactions such as conversation, both sEBR and IEBI values vary considerably for both parties, but until recently this has been considered as something that goes largely unnoticed for both the blinker and their conversational partner [4]. However, recent neuropsychological studies with brain imaging (e.g. [17]) has found evidence that the brain does attend to the blinks of others, even if this is not consciously attended to by the observer. Elsewhere in the literature, there is evidence to suggest that the rate at which people blink affects how others judge them. [19] for example reported that increased blink rates were associated with nervousness and carelessness, while [24] reported that computer generated avatars that blinked at a rate of 18 blinks / min were considered more ‘friendly’, while avatars that blinked at half this rate were judged as being more ‘intelligent’. Studies of changes in blink rate over time continue to generate interesting findings, with several studies for example reporting that people appear to blink less frequently than normal when they attempt to deceive, before then accelerating their blinking rates following the deception (e.g. [16], [18]).

C. Blinking and Robotics

While detailed analysis of the physiological aspects of blinking can be found in the computer graphics research literature (with Disney Research in particular reporting detailed observations of human eye blink movement for use in CGI [26]), literature addressing best practices for the presentation of blinking via robots - particularly the temporal aspects of spontaneous blinking rates - remains somewhat scarce. The most detailed report on temporal blinking behaviours in humans is supplied by Ford et al ([8]) who performed their analysis of blink frequency and duration as a basis for developing an appropriate computational model for use in robotics. Of particular relevance to the present work, Ford notes that some half to two-thirds of all blinking events co-occur with speech events, head movement and mental communicative state changes.

D. Real Time Blink Capture for HRI

To be useful in a tele-presence HRI scenario, it is essential that a blink tracking system has the ability to accurately capture blink events from the teleoperator, and to then be able to modulate them for presentation on the robot in close to real time. Given their relatively long duration (in computational terms) of around 100 milliseconds, capturing eye blink events can be achieved relatively easily using a digital video camera and computer vision-based algorithms, though a reasonably high frame rate is recommended (e.g. [15], [3], [23]). However, real-time detection of blink rates from video does require that the eyes be visible at all times, placing limits on head movement (the camera should ideally be head mounted). Specialist mobile eye trackers offer an alternative approach by bathing the eye with invisible infra-red light, monitoring variances in the reflection of this light as the eyeball moves. However mobile eye trackers are highly expensive, and most trackers that we have worked with do not routinely record the occurrence of spontaneous eye blinks. Though modification of the software to report blink events is feasible, access to the algorithms that handle eye movement data in commercial eye trackers are not generally made available to researchers. Finally, blinks can be captured from the skin surface via electrode, either through electromyograph (EMG) monitoring of the orbicularis oculi muscle of the face (e.g. [6]), or by electroencephalography (EEG, e.g. [25]) which can be measured via the scalp. Calibrated well, both EMG and EEG are extremely accurate, but this requires time, training and skill. Clinical grade recording equipment for both techniques is highly expensive.

For our mobile blink detector, we opted to exploit the ‘Google Glass’ head mounted computing platform. Although not widely reported, the development model of the Google Glass headset contains a small eye-facing IR proximity sensor (see fig. 1) that is designed to detect movements of the right eyelid. The first use of Google Glass to measure blink events was reported in 2014 by [13], who reported a 67% accuracy level across eight participants.

Fig. 1. Proximity sensor embedded in a Google Glass headset (highlighted)

Similar to the system described by [13], in our implementation, a small custom built Google Android application logged the raw values from the proximity sensor continuously. Blink events were detected using a simple peak detection algorithm before being broadcast to our robot via a low latency websocket connection. A Python script installed on the robot responded to incoming blink events by initiating a simple 100msec on-off blink of the LEDS embedded in the
robot’s eyes (see fig: 2). Within the controlled networking conditions used in the study, the delay between the participant’s blink and the corresponding robot blink was very low, appearing to observers as occurring simultaneously.

The Google Glass headset provided us with several notable advantages over the blink capture methods previously described. The Glass headset can operate wirelessly and is - being much lighter than a head mounted eye tracker or camera - relatively unobtrusive. Minimal calibration is required, and the proximity sensor itself is extremely discreet and completely invisible to the participant during its operation. We do however note that the headset that we used is no longer commercially available. However, the IR proximity sensor that it uses is very low cost and widely available. The creation of an analogue using a small microcontroller fitted to a pair of spectacle frames would be straightforward.

II. BLINK TRACKER EVALUATION

A. Method

A controlled laboratory based study was conducted to evaluate the ability of our mobile blink-detector to accurately capture spontaneous eye blinking behaviours, and to transmit / present blink events to a robot in real time under various environmental conditions. To evaluate the performance of our detector under varying degrees of load, we opted to follow the methodology described by [5] and [1], within which spontaneous eye blink behaviours are captured under three activity conditions: 1) at rest [no moving visual stimulus], 2) while reading text from a large display screen and 3) during a reading activity. Video captured during the study was 30FPS AVI, at a resolution of 1920px x 1080px.

B. Participants

28 healthy participants (m=10, f=18) were recruited from staff and students (undergraduate and postgraduate) at the University of Bath. Recruitment was managed through opportunity-based sampling and through advertisement on internal University mailing lists. 45% of our participants were aged between 18 and 24 years old, 41% were between 25 and 34, and the remainder were between 35 and 54. Participants who self reported as experiencing facial ticcing, acute / unusual eye-dryness, or as currently taking medication that they were aware could affect their eye movements were excluded from the study. No other eligibility criteria were applied. Ethical approval for the study was approved by the experimental ethics committee of the University of Bath. Participants were not rewarded for their time.

1) Experimental Setup: Participants completed the task seated at a desk in a quiet and naturally lit laboratory away from direct sunlight. Participants were told that the purpose of the study was to “help us to test the capabilities of a new type of technology that we hope will aid us in our research in robotics”. Participants were informed that they would be expected to wear a Google Glass headset, and that the headset would be tracking their eye movements. However, they were deliberately not told that the Google Glass would be capturing their eye blinking behaviour. This omission was made to minimise our participants becoming overtly conscious of - and potentially disrupting - their natural spontaneous blinking rates.

Throughout the study, a ‘NAO’ humanoid robot (Aldebaran Robotics, model V4) was positioned immediately adjacent to and slightly forward of the participant. This setup allowed the video taping of both participant and robot blinking events within the same shot allowing us to review the footage and establish any problems of network latency and / or packet loss as the headset transmitted blink events to the NAO. The robot was oriented to face away from the participant, remaining still throughout. Participants could not see the robot’s face and they were not aware that the robot was attempting to blink in synchrony with them.

A large format (64”) LCD screen was positioned at eye level, directly in front of the participant at a distance of approximately 2.5 meters. The screen was used to support the filming of the reading section of the study, allowing the participant to complete that task whilst looking straight ahead and thus remaining in camera shot.

The experimenter remained in the room during the evaluation, seated adjacent to the participant but out of their direct line of sight. Video captured during the study was 30FPS AVI, at a resolution of 1920px x 1080px.

C. Procedure

Upon arrival, participants were seated and provided with a set of written instructions, after which their signed consent was obtained. Participants were then instructed to wear the Google Glass headset and to adjust it for comfort. The headset is designed to be wearable without adjustment for head size, but manual adjustment of its prismatic screen display was sometimes required to ensure that its eye facing IR proximity sensor was optimally positioned. To complete the calibration procedure, participants were instructed to adjust the position of the headset’s projected prismatic screen until all four corners were clearly visible, before then turning the screen off. Upon successful calibration, the participant was invited to relax for a few moments before beginning the first task.
To complete the first task, participants were instructed to clear their mind and to sit comfortably, facing straight ahead for a period of three minutes. No talking was permitted. Allowing for a short break, participants were then invited to complete the second task in which the large LCD screen displayed a full-screen ebook version of Lewis Carroll’s *Alice in Wonderland*. The ebook was presented via a web browser using the MagicScroll™ reader application. Font sizes and reading speed were adjusted to allow the participant to read the text comfortably. As with task one, the participants were instructed to sit comfortably in silence and to read the text to themselves for a period of three minutes.

Finally (and again allowing for a short break), in the third and final task the LCD screen was turned off. Participants were instructed that they were now to be interviewed via a series of short questions. The interview section of the study consisted of a short Q&A session of 18 questions, examples of which included ‘If you could have one super power, what would it be?’, ‘When you were younger, what did you want to be when you grew up?’ and ‘what do you do for a hobby outside of studying?’ During the interview, participants were instructed not to look at the interviewer, but to continue looking straight ahead. This was a compromise driven largely by our need to be able to film the face of the participant continually from a fixed position, and to minimise large movements of their head. We accept that a consequence of this approach was the potential loss of some eye-blink events that are known to occur naturally during face to face interaction including (for example) movements of the head and / or large shifts of gaze. The duration of the interview was not fixed, lasting between 2.5 and 3 minutes.

To conclude the study, participants were fully debriefed, and a demonstration of live blink tracking with the NAO robot was provided. The total duration of each session was approximately 20mins.

III. RESULTS

Prior to analysis, all spontaneous blink events were time-stamped manually from the video recordings using the Avidemux software package (version 2.6). 3722 spontaneous blink events were recorded in total. Blink events were extracted by advancing each video recording frame-by-frame (30FPS), with timestamps obtained from the mid-blink frame (the point at which the eyes achieved full closure for a given blink). 'Half blinks', where the eyelids did not completely close, were frequently observed and were included as valid spontaneous blinks. As we were primarily interested in measuring the changes in spontaneous blinking rate levels (i.e. their underlying autonomous blinking rate, as opposed to voluntary blinks that may or may not occur in relation to the blinking behaviour of a interaction partner), blinks that occurred as a result of a sharp turn of the head, or that were clearly voluntary / reflexive in nature were excluded, though these were rare. Such blinks are relatively easy to distinguish, given their longer duration than a spontaneous blink and (in the case of reflexive blinks) visible activation of multiple facial muscles around the eye.

A. Observed Spontaneous Blinking Rates

Exploration of the data revealed a non-normal distribution for blinking rates in all three of our activity conditions. Examination of the Kolmogorov-Smirnov and Shapiro-Wilk tests of normality indicated that both remained highly significant for all conditions. Consequently we chose to employ non-parametric data supportive statistical tests when examining differences in blinking rates between the three activity types. Median spontaneous blink rates per minute were calculated for each of the three activity conditions of *rest, reading* and *interview* and are presented as a plot in fig. 3.

Relative to rest, and following expected population trends, median blink rates were observed to decrease when reading and increase during interview. Median blinking rates / minute for the rest, reading and interview conditions were 13, 7 and 21 blinks per minute, respectively. Comparison of the median blink rates across our sample for each activity using Friedman’s test revealed a statistically significant effect of activity on blinking rate, $\chi^2(2) = 39.297, p < 0.001$.

![Fig. 3. Median blink rates per minute across the three activity types rest, reading and interview](image)

Post-hoc analysis was conducted with Wilcoxon signed-rank tests, with a Bonferroni correction resulting in a significance level of $p < 0.017$. Pairwise comparison revealed a significant difference between all condition pairs: rest and reading ($Z = -3.125, p = 0.002$), rest and interview ($Z = -4.399, p < 0.001$) and reading and interview ($Z = -4.624, p < 0.001$). From this analysis, we were satisfied that our population sample followed known population trends in spontaneous eye blinking rates across these three activity types.

$^1$MagicScroll ebook reader is available at [http://www.magicscroll.net/](http://www.magicscroll.net/)
B. Blink Tracker Accuracy

A summary of the accuracy of the Google Glass blink capture system is presented in Table I. As with the capture of participant blinking events, video records were analysed frame-by-frame. The system failed completely to capture the eye blinks of three of our 28 participants due to hardware and/or network failure. These three participants were therefore excluded from our accuracy analysis.

<table>
<thead>
<tr>
<th>Rest</th>
<th>Reading</th>
<th>Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Blinks</td>
<td>1223</td>
<td>812</td>
</tr>
<tr>
<td>Blinks Captured</td>
<td>1030</td>
<td>683</td>
</tr>
<tr>
<td>Accuracy</td>
<td>84.2%</td>
<td>84.1%</td>
</tr>
</tbody>
</table>

Performance in the less dynamic rest and reading conditions was consistently good, around 84%. Performance in the interview condition was less successful, though we were able to achieve accuracy rates exceeding 80% in over half of our participants. Analysis of the video recordings indicated that the main source of the missed blink events was due to multi-blinks - short rapid bursts of two or three blinks - that occurred too quickly for our current peak detection algorithm to separate. Our analysis found that almost all (26 of 28) participants exhibited some degree of multi-blink behaviour, with a total of 364 multiblinks being recorded within the total count of 3722 (9.78%). Multi-blinks occurred more frequently in the interview condition, but the distribution of double blinks was not consistent across our participant pool. For the majority (22 participants) the occurrence of multiblinks was sporadic and infrequent, making up less than 10% of their total spontaneous blinks. However, the remaining eight participants displayed much higher multiblink rates of between 15% and 26% of their total blinks.

C. Changes in Blink rate when Speaking and Listening

During our analysis of the video recordings, we consistently observed a noticeable difference in blinking rates when a participant was listening compared to when they were speaking. To explore this phenomena further, separate counts of blinks that occurred when listening and responding were calculated for each participant. Given that the duration of each question and answer during the interview varied, a blink rate / minute value for each question was calculated separately to allow the data to be comparable across participants. As with our analysis of blinking rates across condition, the distribution of the data was again found to be non-normal. Median spontaneous blinking rates / minute whilst speaking and listening in the interview condition are presented in Fig. 4.

A Wilcoxon Signed-Ranks Test indicated that the median blinking rates observed whilst listening (Mdn = 53) were higher across our sample than the median blinking rates observed whilst responding (Mdn = 35). This difference was statistically significant at the .001 level, Z = -3.605, p < 0.001. The median difference was -18 blinks per minute, indicating that people blink roughly 33% more frequently when they are listening compared to when they are speaking.

IV. DISCUSSION

In this paper, we presented an evaluation of a lightweight mobile spontaneous eye blink capture system with 28 participants, re-presenting their blink events in near real time via an Aldebaran NAO model robot. We argue that, given the established link between eye blinking rates and cognition, presenting the eye blinking behaviours of a tele-presence robot pilot via their proxy could offer a useful low-bandwidth facial cue for tele-presence HRI that is worthy of more detailed investigation.

Comparison of the blinking behaviours of our participant sample to known population trends indicated that the performance of our sample was within expected ranges. Across three activity types, the current instantiation of our blink tracker system performed with a general accuracy rate of around 80%, comfortably exceeding the accuracy rates reported by [13] but with clear room for improvement.

Though our tracker currently performs well under consistent light load (i.e. at rest and while reading), it struggled within the much more dynamic interview activity used in our evaluation. Two immediate sources of inaccuracy have been identified: difficulty in capturing 'half blinks' (where the eyelids do not completely close) and failure to capture rapid multi-blinks. Both suggest that our current peak detection algorithm was too conservative and that by tuning the threshold and efficiency of our algorithm, we should be able to increase the accuracy of our system substantially. Work towards this is ongoing.

Further to our evaluation of the accuracy of our tracking system, we also noted an observation of blinking behaviour that would be useful to other researchers in this area.
Blinking rates across our sample increased by around 33% when our participants were listening compared to when they were speaking. To our knowledge, this finding is new to the literature, though more work is required before we are able to generalise to wider populations with confidence. However, if this is shown to be the case, we suggest that this change in rates could be useful in general robotics interaction design as a method for indicating a robot’s state (e.g. paying attention, listening etc).

A. Limitations and Future Work

There are several limitations of our current blink detector that are the focus of future research. A significant limitation of our detector is that it currently only measures blink events (via peak detection), and not individual blink durations. Capturing blink duration would be immediately useful to help differentiate between spontaneous and voluntary / reflexive blinks and may also have social communicative value that we have yet to consider. Modifications to our detection algorithm to capture duration are currently being explored.

A second limitation of our evaluation is that, while the wireless web-socket connection used to transfer the blink events to our lab robot performed well over a local network, we currently have no data on performance in larger and less controlled network environments. We also note that the development model of the Google glass headset that we used can overheat, resulting in problems with the on-board WiFi.

Finally, although our population sample followed expected behavioural trends, and though we made efforts to exclude volunteers with obvious blink rate abnormality, we did not consider the impact of current cognitive state. Fatigue levels and presence of stimulants (e.g. caffeine, nicotine) for example were not controlled for, and the impact of these factors is currently unknown.

ACKNOWLEDGMENT

The authors thank Ms. Phaedra Florou and Ms. Lauren Barnes for their support during the study.

REFERENCES