Int. J. Human-Computer Studies 92-93 (2016) 55–65

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Gaze perception and awareness in smart devices

Muhammad Sikandar Lal Khan [b], Haibo Li [a], Shafiq ur Réhman [b,*]

[a] CSC-Media Technology and Interaction Design, Royal Institute of Technology (KTH), Stockholm, Sweden
[b] Immersive Interaction Lab (i2lab), Department of Applied Physics and Electronics (TFE), Umeå University, 90187 Umeå, Sweden

ARTICLE INFO

Article history:
Received 10 February 2015
Received in revised form
10 March 2016
Accepted 6 May 2016
Available online 9 May 2016

Keywords:
Mona-Lisa Gaze Effect
Gaze awareness
Computer-mediated communication
Eye contact
Head gesture
Gaze faithfulness
Embodied telepresence system
Tablet PC
HCI

ABSTRACT

Eye contact and gaze awareness play a significant role for conveying emotions and intentions during face-to-face conversation. Humans can perceive each other's gaze quite naturally and accurately. However, the gaze awareness/perception are ambiguous during video teleconferencing performed by computer-based devices (such as laptops, tablet, and smart-phones). The reasons for this ambiguity are the (i) camera position relative to the screen and (ii) 2D rendition of 3D human face i.e., the 2D screen is unable to deliver an accurate gaze during video teleconferencing. To solve this problem, researchers have proposed different hardware setups with complex software algorithms. The most recent solution for accurate gaze perception employs 3D interfaces, such as 3D screens and 3D face-masks. However, today commonly used video teleconferencing devices are smart devices with 2D screens. Therefore, there is a need to improve gaze awareness/perception in these smart devices. In this work, we have revisited the question; how to improve a remote user's gaze awareness among his/her collaborators. Our hypothesis is that '*an accurate gaze perception can be achieved by the '3D embodiment' of a remote user's head gesture during video teleconferencing*'. We have prototyped an embodied telepresence system (ETS) for the 3D embodiment of a remote user's head. Our ETS is based on a 3-DOF neck robot with a mounted smart device (tablet PC). The electromechanical platform in combination with a smart device is a novel setup that is used for studying gaze awareness/perception in 2D screen-based smart devices during video teleconferencing. Two important gaze-related issues are considered in this work; namely (i) 'Mona-Lisa Gaze Effect' – the gaze is always directed at the person independent of his position in the room, and (ii) 'Gaze Awareness/Faithfulness' – the ability to perceive an accurate spatial relationship between the observing person and the object by an actor. Our results confirm that the 3D embodiment of a remote user head not only mitigates the Mona Lisa gaze effect but also supports three levels of gaze faithfulness, hence, accurately projecting the human gaze in distant space.

1. Introduction

An estimated 60–80 percent of human communication is facilitated by non-verbal communications, e.g., facial expression, head gesture, eye-contact, gaze direction, and body language (Mehrabian, 1977). These non-verbal cues together with verbal information give meaning to our conversation and facilitate us in making decisions. Among all these non-verbal cues, gaze direction plays a central role for conveying information among human beings (Kendon, 1967). The significance of gaze in face-to-face communication has been widely studied (Yarbus, 1967; Kleinke, 1986). These studies show that the gaze is one of the important non-verbal cues in face-to-face communication (Langton et al., 2000), with 60 percent of communication involves gaze and 30 percent of communication involves mutual gaze (Kendon, 1967; Peters et al., 2005). Gaze implies the 'direction of one's eyes towards a person or an object'. The gaze is different from eye contact or mutual gaze, which implies that the 'two persons are looking at each other simultaneously'. Human estimate the gaze direction of a person by the combined information of eyes, head and body movements.

The gaze behaviours of humans make a substantial contribution in interpersonal and social interactions (Argyle and Cook, 1976). These gaze behaviours are used for regulating the flow of conversation by turn taking, side conversation, synchronised verbal and kinaesthetic behaviours (Rutter and Stephenson, 1977). The research studies show that these gaze behaviours influence the evaluation of the attentiveness, competence, dominance, credibility, social skills, intensity of their feelings, intimacy, etc. (Thayer and Schiff, 1977). Furthermore, researchers have also

* Corresponding author.
E-mail addresses: muhammad.sikandar.lal.khan@umu.se (M.S.L. Khan), shafiq.urrehman@umu.se (S. ur Réhman).

http://dx.doi.org/10.1016/j.ijhcs.2016.05.002
1071-5819/

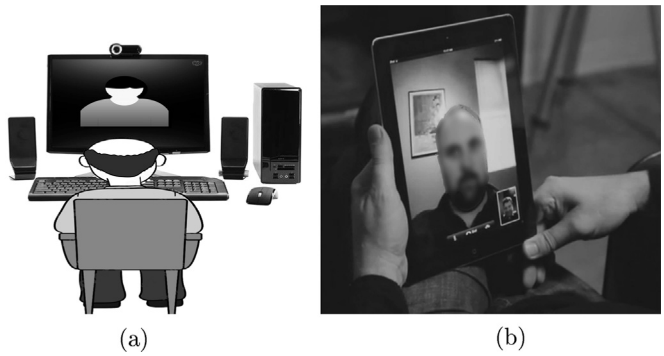

Fig. 1. Standard video conferencing setups, (a) a desktop computer with a webcam mounted on top of the LCD/LED screen and (b) in case of a smart device, a person holds an iPad for video conversation.

investigated the function of gaze in social control, such as, persuasion and deception, threat and dominance, escape and avoidance, and ingratiation and compliance (Kleinke, 1986). The findings of these studies advocate a high significance of gaze in human–human communication.

The significance of gaze in human–human communication compels researchers to study gaze behaviours in computer-mediated communication, e.g. video teleconferencing. Video teleconferencing occurs when a small group of geographically distantly located people holds discussions in real time, during which they hear and see each other through computer-based devices. The research studies show that the distantly located people cannot estimate the gaze and cannot even make an eye contact during traditional video teleconferencing (Chen, 2002; Mukawa et al., 2005; Bekkering and Shim, 2006). Traditional video conferencing systems rely on a camera displaced relative to the image of a remote user, which leads to an immediate misalignment and displaced gaze direction (see Fig. 1). The second major reason is the use of 2D flat displays for video teleconferencing. The 2D rendition of a 3D human face poses a problem for preserving the correct gaze direction.

Today, smart devices are driving the technology of daily life. There are more smart devices than human beings on the planet, and these devices are multiplying five times faster than the human population (Cisco System). It is expected that these smart devices could increase to 10 billion by the end of year 2016 (Cisco Systems). These smart devices are becoming a vital source of communication ranging from text, voice and video communication even for users with disabilities (ur Réhman and Liu, 2010; ur Réhman et al., 2014). By the constant increase in the network speed and the computational power of multimedia-capable cellular phones, video communication is becoming increasingly popular.

In this work, we have revisited the problem of 'gaze perception and awareness' during video teleconferencing and propose a novel solution based on the concept of '3D embodiment' of a remote person. The 3D embodiment is achieved by our embodied telepresence system (ETS), which uses commonly-used smart device with 3-DOF electromechanical platform. This setup presents the head gestures of a remote person along with standard audio-video communication. The ETS was previously used for embodied interaction-based video teleconferencing, where head gestures and other social functions are considered (Khan et al., 2014a,b). In this work, we have performed a *gaze-related* experiment with ETS to test our hypothesis, 'an accurate gaze perception can be achieved in distance communication by the '*3D embodiment*' of a remote person'. Our experiment considers two discrete gaze-related issues:

- Mona Lisa gaze effect.
- Gaze Faithfulness/awareness.

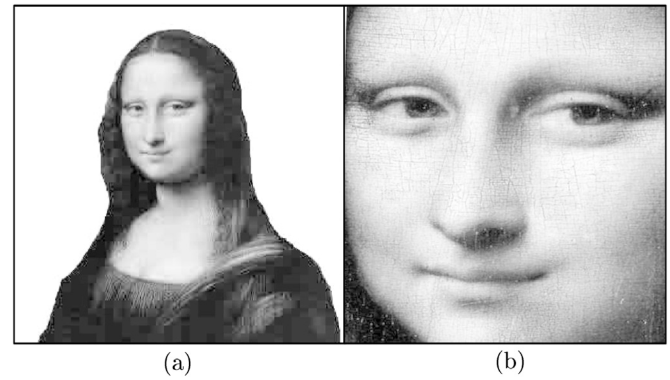

Fig. 2. Mona Lisa Gaze Effect: from left to right: the original image of Mona Lisa and the zoomed image of Mona Lisa. From these images, it is clear that Mona Lisa is looking directly at you irrespective of your position. (This picture is in the public domain).

The Mona Lisa gaze effect is when Mona Lisa's gaze is always directed to you when you move in any direction in the room (see Fig. 2) (Boyarskaya and Hecht, 2009). It is argued that this effect is observed when a remote person-face/picture is embedded in a virtual 2D screen/frame and the gaze of a remote person does not specify any target person or an object. The Mona Lisa gaze model for video teleconferencing has recently been investigated by Al Moubayed et al. (2012). This gaze model consists of several assumptions; two of these assumptions are of particular interest for our experiment.

1. The Mona Lisa gaze effect is more general and does not rely on just human eyes and head direction, as illustrated in Fig. 3.
2. For accurate gaze estimation, the angle of the eyes is not sufficient, an estimation of the position and angle of the head is also required as illustrated in Fig. 4.

Al Moubayed et al. (2012) have proposed a 3D face mask to mitigate a Mona-Lisa gaze effect in video teleconferencing. However, we have considered a 2D display with 3D embodiment of human head to mitigate the Mona Lisa gaze effect. Our experimental study extends investigations to see whether our proposed system supports gaze faithfulness/awareness. The 'Gaze Awareness/Faithfulness' is related to gaze direction information. It is an ability to perceive an accurate spatial relationship between an observing person and the object, that is being observed (Monk and Gale, 2002). There are three types of gaze awareness that are of most interest in a video teleconferencing (Monk and Gale, 2002);

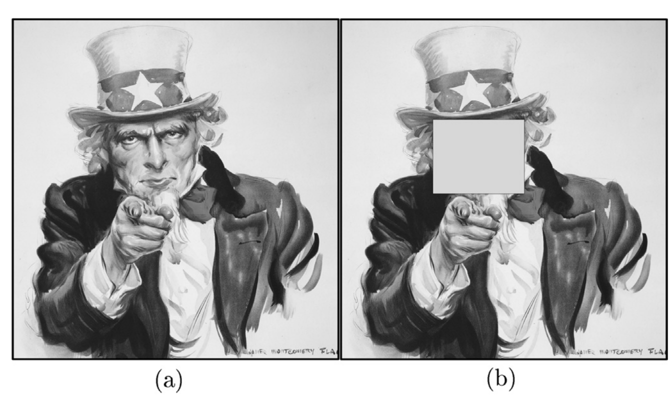

Fig. 3. I want you for the US Army: from left to right: (a) an image of Uncle Sam; (b) faceless image of Uncle Sam. It is clear that Complete Uncle Sam and Faceless Uncle Sam are pointing at you. (This picture is in the public domain and taken from Al Moubayed et al., 2012).