Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 25 (2014) 205 – 212

www.elsevier.com/locate/procedia

8th International Conference on Digital Enterprise Technology - DET 2014 – "Disruptive Innovation in Manufacturing Engineering towards the 4th Industrial Revolution"

Integrating Pointing Gesture Detection for enhancing Brainstorming Meetings using Kinect and PixelSense

Andreas Kunz[a]*, Ali Alavi[a], Philipp Sinn[a]

[a]*Innovation Center Virtual Reality (ICVR), ETH Zurich, 8092 Zurich, Switzerland*

* Corresponding author. Tel.: +41-44-63-25771 ; fax: +41-44-63-21125. *E-mail address:* kunz@iwf.mavt.ethz.ch

Abstract

Microsoft PixelSense is able to detect multitouch input and tagged objects as well, which makes it suitable to be used in net-based brainstorming sessions within small teams. However, any gestures above the table cannot be detected, which makes net-based brainstorming sessions less intuitive. Thus, we present a solution how Kinect can be used together with PixelSense to overcome this limitation without interference between the two devices.

Peer-review under responsibility of The International Scientific Committee of the 8th International Conference on Digital Enterprise Technology - DET 2014 – "Disruptive Innovation in Manufacturing Engineering towards the 4th Industrial Revolution"

Keywords: Kinect; Deictic Gestures; Computer Supported Collaborative Work

1. Introduction

Nowadays, many big ideas incept from a team of individual designers, sitting together in a brainstorming meeting. This collaborative approach uses the collective knowledge and creativity of the team, and is beneficial to the major stakeholders of the meeting (the designers, the organization, and the customers) [1]. The advent of computers and digital media, and increased availability of high speed networks, enabled remote brainstorming sessions. Thus, members of a group, distributed among different locations, may participate in such meetings and effectively participate in the ideation phase of developing a product. This has two main advantages:

Firstly, the need of commuting between different locations can be eliminated. This is particularly interesting for international firms with offices located around the globe.

Secondly, there are reduced social inhibitions among group members. It has been seen that many team members, particularly in the presence of a senior member, will withhold commenting for fear of criticism or negative evaluation, a behavior that depresses the ideative efficacy [2]. Moreover, in Electronic Brainstorming Systems (EBS), "because participants do not see each other (even if they are in the same room), attention is essentially paid to ideas […] helping to reduce redundancy [of ideas] and improve task performance" [3].

In spite of these advantages, net-based brainstorming has certain disadvantages, mainly due to the fact that available EBSs are not capable of transmitting different forms of interaction: as depicted in Figure 1, these interactions happen between two types of entities, and in two different spaces (see Figure 1):

Interaction between humans and the digital media happens on the "task space", which can be a tabletop computer, and above it, for example pointing to the contents on the table. This is where the generated artifacts (forms of mindmaps, sketches or written notes) belong to.

Interaction between team members, which according to [3] consists of verbal (words), vocal (intonation), and visual (body language) elements, takes place in communication space.

For an efficient net-based brainstorming, we should decide on a subset of these communication elements to be captured, aligned, and correctly transferred to the remote side.

While most EBSs are capable of transmitting the content of the task space and also verbal and vocal elements of the communication space, they come short in a proper

Peer-review under responsibility of The International Scientific Committee of the 8th International Conference on Digital Enterprise Technology - DET 2014 – "Disruptive Innovation in Manufacturing Engineering towards the 4th Industrial Revolution"
doi:10.1016/j.procir.2014.10.031

transmission of visual parts of human-to-human and human-to-content communication (facial expressions, hand gestures, nodding, shrugging, and so forth).

Many approaches try to overcome this problem by delivering this visual content using video-conferencing, which is not efficient. Firstly, many researches in social psychology show that lack of social cues in brainstorming meetings leads the group members to focus on the task instead of on the people, thus improving the task performance [3], [5]. In other words, delivering all the social content of the meeting is not ideal. As a result, video-conferencing may increase social inhibition among group members. Moreover, in order to view remote collaborators, a big portion of the digital media screen needs to be dedicated to showing other collaborators faces and bodies, leading to either a very small task space, or the need of additional screens. Both adversely affect the quality of the meeting. Hence, we are interested in transmitting only the essential visual communication elements which contribute to the meeting while not adding much social inhibition. Additionally, it does not occupy significant visual resources of the meeting or distracts the participants' attention.

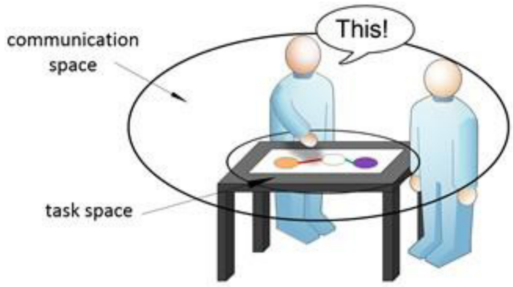

Fig. 1. Brainstorming integrates communication- and task-space.

In the following paper, we introduce a new technology for detecting pointing gestures, relating them to the content of the task, and informing the remote partners about it. The main part of this process is to detect and track the pointing gestures. Since pointing gestures are typically performed in the free space (communication space) above the interactive table, PixelSense's sensors cannot detect them anymore with a sufficient resolution. Thus, an additional tracking system is required. Although a marker-based tracking system can accomplish this task, we are interested in a solution that is less intrusive to the user, because wearing the markers might neither be possible nor desirable for some of the participants. Moreover, detecting hands and fingers with color-based cameras is not an option neither, because of the variety of skin tones in a large group of people, and also a large range of colors available in the task space. Hence, we try a different approach by tracking hands using a depth camera. Even though tracking hands using depth data does not have the mentioned problems of other tracking systems, a depth camera cannot be easily used in presence of a PixelSense touch screen, since they both work with infrared light, thus interfere with each other. This paper offers a solution to overcome this interference problem between the Microsoft Kinect depth camera and the Microsoft PixelSense touch screen. Once the pointing gesture's orientation is detected, the remainder of the paper will describe, how the target of the pointing gesture will be displayed to the remote partners.

2. Related Work

The importance of aligning the layers from Figure 1 was already stated in earlier work. According to Ishii et al. [6], people feel it difficult to communicate if they cannot tell whether the remote partner is listening carefully or not. An early example of a shared workspace was given by Krueger [7]. However, in his setup the shared workspace was more a shared view space, since it was not possible to interact with the artifacts. In 1990, Tang and Minneman [8] introduced VideoDraw, a device that allows the partners to share a drawing surface. It consists of video cameras aiming at the screen, whereby each camera is connected to a monitor on the other side. As both partners draw with whiteboard markers on the screen, the video camera captures these markers and the accompanying hand gestures, which are then transferred to the other side. VideoWhiteboard [9] features rear-projection of the shared task space, and a camera which is also placed behind a 90'' projection screen. The partners see the complete image, real and video marks, as well as the shadows of their remote partners' gestures and actions. Bly [10] conducted an exploratory study to investigate the use of a drawing surface in design sessions. In one of her settings, two designers were geographically separated and connected via video tools. Bly observed that in the sessions that provided visual contacts, "gestures constituted a significant portion of the drawing actions that took place". In order to allow designers to work remotely by sharing a drawing surface, Bly and Minneman developed Commune [11]. This system provides two separate horizontal writing surfaces, each consisting of a horizontally mounted CRT monitor, and a transparent digitizing tablet mounted directly on top of the screen. On each writing surface, collaborators can gesture and make marks by using a stylus. However, the remote partner was not captured, but his gestures were restricted to a telepointer that was transferred to the remote site.

With ClearBoard [6], Ishii et al. bring together task space and communication space, since the system allows keeping eye contact while working on an interactive surface. Kirk et al. [12] [13] underlined in their study the importance of hand gestures that are in correct relation to the task space. Stotts et al. [14] suggested using remote collaboration groupware that displays live video embodiments situated within the shared workspace. With "The Vis-a-Vid (VAV) Transparent Video Facetop", they presented a respective user interface. It has cameras that acquire live video embodiments showing the collaborators' faces. The local live video embodiment is displayed as visual feedback for controlling where the hand is placed. Further, pointing gestures are detected to allow controlling the computer's mouse pointer. Due to the camera positions, gestures must be performed in free space, making VAV less useful for on-screen interaction.

Wellner presented the Digital Desk [15] and the Double Digital Desk [16]. The Digital Desk consists of a normal office desk with a projector and a camera above it, both pointing to the desk's surface. The projector allows