Future Generation Computer Systems 54 (2016) 313–325

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Toward natural multi-user interaction in advanced collaborative display environments

Vít Rusňák [a,*], Lukáš Ručka [a], Petr Holub [a,b]

[a] Masaryk University, Botanicka 68a, Brno, Czech republic
[b] CESNET z.s.p.o., Zikova 4, Praha, Czech Republic

HIGHLIGHTS

- The concept of cue-less multi-user interaction.
- An approach for building user-aware interactive environment.
- Combination of multiple modalities to enable association of input events with users.
- Interaction framework for group collaborative environments.
- Building multi-sensor environments from commodity hardware.

ARTICLE INFO

Article history:
Received 31 May 2014
Received in revised form
12 January 2015
Accepted 24 March 2015
Available online 1 April 2015

Keywords:
Group collaborative environments
Multi-user interaction
Scalable resolution displays
Human factors
Distributed systems

ABSTRACT

Large high-resolution displays have become widely spread in research centers, laboratories, and public spaces during the last decade. There have been various research efforts in transformation of these displays from passive screens to interactive environments where multiple users can interact independently with multi-touch surfaces, in-air interaction or pointing devices. Such systems provide basic multi-user interaction, i.e., two or more users can interact concurrently. However, continuous user tracking and association of input events with users, which could considerably improve user experience, is still a largely unexplored topic. In this paper, we present a set of techniques enabling cue-less multi-user interaction in environments where horizontal or vertical high-resolution displays are in the role of central visualization platforms. We define the concept of cue-less multi-user interaction and set of techniques that enable unobtrusive user tracking and their association with input events through combination of a multi-touch surface and a depth sensor. Further, we introduce an open-source framework implementing techniques behind this concept and their evaluation in terms of accuracy of user association and the overall system interactivity when integrated with selected distributed rendering middleware.

1. Introduction

The rapid expansion of computer systems into our environment makes them ubiquitous. From smartphones, through tablets to large screens, which occupy research laboratories, public spaces, and our homes, one can observe paradigm shift toward interaction with them. Over three decades, keyboard and mouse were considered as the primary interaction devices for computer users. Natural user interfaces (NUIs) are seamless interfaces, where users interacts with the system directly with their fingers, palms or other body parts, or generally without using artificial control devices. NUIs are broadly explored and they become one of the main directions in the human–computer interaction research.

With touch-responsive surfaces, displays become responsive to users' input, which creates new possibilities for their integration, for example, within group collaborative environments. Single-display groupware tools are represented by tabletops or large display walls operated by several users [1]. The concept of digital control rooms or war room combine multi-modal interaction with multiple of these displays and other devices to make the use of technology efficient and intuitive [2–4]. Growing demands on visualization of complex data caused worldwide expansion of tiled-display systems to the hundreds of laboratories and research centers, where they provide an efficient tool for visual analysis as well as for presentations and demonstrations [5,6].

* Corresponding author.
E-mail addresses: xrusnak@fi.muni.cz (V. Rusňák), xrucka@fi.muni.cz (L. Ručka), hopet@cesnet.cz (P. Holub).

http://dx.doi.org/10.1016/j.future.2015.03.019
0167-739X/

Table 1
Comparison of interaction types in zones.

Interaction zones	Tabletop			Display Wall		
	Direct		Indirect	Direct		Indirect
	Touch	Touch-less		Touch	Touch-less	
Up-close	(all)		Mouse + keyboard [12] widget [13]	[1,8]	[14]	
Mid-air		[15]			[14]	Game controller [1], smartphone [16–19]
Remote						PC [10]

The integration of multi-touch interfaces to tiled-display rendering platforms [7,8] enabled basic multi-user interaction, i.e., multiple users can interact concurrently with the environment. However, basic multi-user interaction approach is incapable of distinguishing individual users, which leads to conflicts in collaborative tasks that regularly happen. For instance, when two users want to reposition an application window in opposite directions at the same time, the trajectories of users' fingers are in the opposite direction similar to spread gesture. Thus in a basic multi-user interaction, the window is actually magnified. Instead, the correct response of the system should be repositioning according to the finger trajectory of the user who touched the surface earlier. Implementing association of users with input events within these systems transforms them into complex group collaborative environments [6] where the interaction with the system become ubiquitous.

We define *cue-less multi-user interaction* as the interaction paradigm that extends the basic multi-user interaction, in which the group collaborative system can differentiate individual users and associate them with their input events. Users are free to move around within the interaction area, being tracked with no artificial markers nor any personal tracking device. Cue-less multi-user interaction requires complex description of static components and dynamic behavior of the system. Static component description defines the interaction space given by placement of physical sensor devices. Distinguishing and tracking of users, their body parts and processing of input (touch) events are the dynamic components of the system.

Our work is motivated by improving interaction with digital content (e.g., videos, images, visualizations) by metaphors such as possession of digital objects or inherent user distinguishing. We focus on the overall system unobtrusiveness and try to achieve walk-up-and-use user experience; i.e., first-time or one-time user can use the system without *a priori* orchestration and training.

In this paper, we explore how we can improve multi-user tactile interaction by associating users with input events in group collaborative environments, based on horizontal and vertical tiled-display systems. We have prototyped our approach using OptIPortals [9], a state-of-the-art group collaborative environment, which combines cluster computing performance and low-latency photonic networks to ensure high-level of interactivity.

Specifically, we make the following contributions to the state-of-the-art: (a) we propose novel techniques that enable cue-less multi-user interaction in an unobtrusive manner in these systems (i.e., marker-less and without wearables) and scalable approach to building multi-touch surfaces for tiled-display systems; (b) we describe the prototype of Multi-Sensor (MUSE) framework that implements the above techniques and we discuss important considerations for such a system; (c) we investigate the performance and accuracy of the proposed techniques implemented within the MUSE framework. The MUSE framework installation has become a part of the SAGE-based [10,11] OptIPortals installation in the Laboratory of Advanced Networking Technologies (SITOLA), which is part of the CineGrid and GLIF infrastructures.

The rest of the paper is organized as follows. Section 2 discusses the state of the art and related work in multi-user interaction techniques and user distinguishing approaches in tabletop and display wall environments. Techniques enabling cue-less multi-user interaction and proposed novel algorithms are discussed in detail in Section 3. Section 4 describes the MUSE framework implementing key concepts of cue-less multi-user interaction. Evaluation of the prototype implementation is provided in Section 5 and perspective of our future work follows in Section 6.

2. Related work

The interaction with computer systems is either direct (touch-based or voice-controlled) or indirect with pointing devices (game controllers, mouse/gyro-mouse) and/or keyboards. Unlike desktop computers where the user sits in front of the display, there can be more *interaction zones* in larger systems, such as room-sized group collaborative environments [7]. The zones are differentiated based on the distance from the interactive surface (or displays), positioning accuracy, or comfort of the users. We distinguish three interaction zones-up-close, mid-air (or distance) and remote. Remote interaction is usually implemented as a single-user graphical or textual control interface running on a separate device and is out of scope of this paper. Particular attention is paid to research problems in natural multi-user interaction within up-close and mid-air interaction zones: e.g., continuous user tracking and association of users with their input events (or commands), or identification of users and their roles.

2.1. Multi-user interaction with display systems

This section provides an overview of the state of the art of the up-close and mid-air interaction techniques. Summary of zones and interaction types summarizes Table 1.

Up-close interaction zone. This zone is suitable for very accurate operations and users typically interact directly with the surface. Up-close indirect interaction is specific for tabletops only, where various tangible objects or keyboards with mice could be used [20,13]. Frustrated Total Internal Reflection (FTIR) and Diffused Illumination (DI) are commonly used multi-touch sensing technologies, but they are limited to rear-projected systems such as tabletops and they cannot be used with systems based on common display technologies such as LCD and OLED displays.

An alternative approach is to use multi-touch frames with IR emitters and sensors [1,8]. Although such frames were developed for use with display walls, the number of concurrent touches is still limited (up to 32 in most cases) as well as their scalability in terms of the overall size of the interactive surface. Touch-less interface for direct interaction constructed as an array of RGB cameras in front of display wall has been proposed by Stødle et al. [21]. Form-factor capacitive overlay sensors used in smartphones, tablets and displays with multi-touch surfaces, are manufactured in relatively small sizes (up to 24″) providing approximately up to 16 concurrent touch points. Current multi-touch sensors are capable of recognizing multiple concurrent touch events, but association of the multiple events must be done using some additional means.

Mid-air interaction zone. Interaction with tabletops and display walls from a medium distance is very effective for layout and global positioning of the displayed content. Users interacting from the