MAUI: Tele-Assistance for Maintenance of Cyber-Physical Systems

Philipp Fleck[1] , Fernando Reyes-Aviles[1] , Christian Pirchheim[1] ,
Clemens Arth[1,2] and Dieter Schmalstieg[1]
[1]*ICG, Graz University of Technology, Inffeldgasse 16/2, 8010 Graz, Austria*
[2]*AR4 GmbH, Strauchergasse 13, 8020 Graz, Austria*
{*philipp.fleck, fernando.reyes-aviles, pirchheim, dieter, arth*}*@icg.tugraz.at*

Keywords: Remote Collaboration, Telepresence, Augmented Reality

Abstract: In this paper, we present the maintenance assistance user interface (MAUI), a novel approach for providing tele-assistance to a worker charged with maintenance of a cyber-physical system. Such a system comprises both physical and digital interfaces, making it challenging for a worker to understand the required steps and to assess work progress. A remote expert can access the digital interfaces and provide the worker with timely information and advice in an augmented reality display. The remote expert has full control over the user interface of the worker in a manner comparable to remote desktop systems. The worker needs to perform all physical operations and retrieve physical information, such as reading physical labels or meters. Thus, worker and remote expert collaborate not only via shared audio, video or pointing, but also share control of the digital interface presented in the augmented reality space. We report results on two studies: The first study evaluates the benefits of our system against a condition with the same cyber-physical interface, but without tele-assistance. Results indicate significant benefits concerning speed, cognitive load and subjective comfort of the worker. The second study explores how interface designers use our system, leading to initial design guidelines for tele-presence interfaces like ours.

1 INTRODUCTION

Technical facilities increasingly incorporate cyber-physical systems, which must be operated and maintained using a mixture of physical and digital interfaces. This evolution is not limited to industrial settings: For example, many components of a modern car can no longer be repaired with physical tools without access to diagnostic devices that access the car's software-controlled sensors and actuators. In industrial environments, workers must learn to operate a mixture of conventional physical controls, such as levers or buttons, and digital interfaces, such as touch panels placed next to machines or desktop computer interfaces in a control station. For maintenance and repair, workers must access sensor telemetry data through diagnostic interfaces, while cross-referencing the telemetry with sources on paper. Mastering this multitude of information sources can be difficult. For cases where the local worker's competence is exceeded, equipment manufacturers operate call-centers or send experts to customers at high traveling costs.

Consequently, tele-assistance has been proposed as a relief for workers and experts. Most tele-assistance solutions establish a shared presence of remote expert and worker via some form of audio/video link combined with tele-pointing and on-screen annotation. One approach is to combine tele-presence with augmented reality (AR) by letting the worker wear a head-mounted display (HMD) with built-in camera and microphone. The worker has their hands free and does not have to switch attention to a stationary screen in order to access digital data. The remote expert can see what the worker sees, provide spoken instructions, and point out relevant areas in the shared video.

These collaborations primarily target supporting the mutual spatial understanding of worker and expert. They are sufficient in situations where only physical interfaces or facilities need to be manipulated. However, in a cyber-physical system, digital interfaces must be manipulated as well, requiring shared understanding of non-spatial aspects. Collaboration in digital space is common practice for office workers, who combine video conferencing with screen sharing and remote desktop interfaces. The expert will commonly operate the office worker's desktop computer remotely. In such a tele-assistance situation, the physical interface (*e.g.,* , feeding paper into a printer) must be operated by the worker, while the digital interface is subject to mixed initiative (both users take turns at operating the mouse). Oftentimes the digital interface will almost exclusively be operated by the expert, in particular, if the worker lacks technical understanding.

In this work, we combine AR tele-assistance with shared operations via a remote desktop-like user interface into a framework called *maintenance assistance user interface* (*MAUI*). Our system is able to display web-based interfaces as overlays in the HMD. These

Figure 1: Facility for production. (left) An expert in the control center dispatching (right) a worker on the shop floor, wearing a HoloLens, to fix machine downtime problems.

digital interfaces can contain any mixture of telemetry information, instructions and interactive widgets used to control the digital aspects of a cyber-physical system. Since the worker is busy with the physical tasks, may lack experience with the digital interface, or is simply overwhelmed by the amount of information displayed in the relatively small field of view of the HMD, the primary responsibility for configuring and operating the digital interface is deferred to the remote expert. The expert helps the worker to interact with the virtual and real world at the same time, thus reducing the risk of possible failures or mistakes. MAUI was designed with the intention of more efficient subdivision of work between expert and worker, providing the following contributions:

- We analyze the requirements of AR tele-assistance for industrial jobs, where remote operation and configuration of a user interface by a remote expert is required.
- We present implementation details of MAUI, which address the aforementioned requirements with its robust abilities for sharing audio, video, digital content and control state in harsh industrial environments.
- We discuss results of a user study demonstrating lower task completion times, reduced cognitive load, and better subjective comfort on performing a maintenance task.
- We discuss the results of an exploratory study, orthogonal to the first one, analyzing how web developers perform when creating user interface content in the MAUI framework.

2 RELATED WORK

AR can help a worker by purely displaying digital information. Henderson *et al.* (Henderson and Feiner, 2011) have shown the benefits of AR in the maintenance and repair domain. However, such pre-configured information sources are often unavailable. In this case, a good alternative is to link the worker to a remote expert providing live support. Dealing with a cyber-physical system adds the dimension of a digital interface, which can be controlled locally by the worker, or alternatively by the remote expert. Thus, our work is at the nexus of remote collaboration, interaction with cyber-physical systems and remote desktop user interfaces. We provide background to each of these topics in the remainder of this section.

2.1 Remote collaboration

Video transmission has been the enabling technology for tele-assistance, since the pioneering work of Kruger (Krueger et al., 1985). Early work in this space (Barakonyi et al., 2004; Wellner and Freemann, 1993) was mostly constrained to desktop computers, due to technical limitations. Recent progress in mobile and wearable computing has brought video conferences abilities to the factory floor.

However, establishing a shared spatial presence at the task location still proves challenging. Experts need to visually experience the worker's environment. The video stream from a camera worn by the worker will only show what the worker is seeing (Huang and Alem, 2013; Bauer et al., 1999; Kim et al., 2013; Chastine et al., 2008; Kurata et al., 2004). Giving the remote expert independent control of a robotic camera (Kuzuoka et al., 2000; Gurevich et al., 2012) is usually not economically feasible.

Apart from spoken instructions, most tele-assistance solutions let the expert provide visual-spatial references, either via hand gestures (Oda et al., 2013; Huang and Alem, 2013; Kirk and Stanton Fraser, 2006), remote pointing (Bauer et al., 1999; Fussell et al., 2004; Kim et al., 2013; Chastine et al., 2008), or hand-drawn annotation on the video (Fussell et al., 2004; Gurevich et al., 2012; Chen et al., 2013; Kim et al., 2013; Ou et al., 2003). Hand-drawing is either restricted to images from a stationary camera viewpoint (Kim et al., 2013; Gurevich et al., 2012; Huang and Alem, 2013; Bauer et al., 1999; Chen et al., 2013; Fussell et al., 2004; Kirk and Stanton Fraser, 2006; Kurata et al., 2004) or requires real-time 3D reconstruction and tracking (Chastine et al., 2008; Lee and Hollerer, 2006; Gauglitz et al.,