Conferinta Nationala de Interactiune Om-Calculator – RoCHI 2014

On the Enhancement of Augmented Reality-based Tele-Collaboration with Affective Computing Technology

Dragoş Datcu
Faculty of Technology, Policy and Management, Delft University of Technology
Jaffalaan 5, Delft, 2628BX The Netherlands
D.Datcu@tudelft.nl

ABSTRACT

The paper describes the integration of two major fields of research, namely tele-collaboration by augmented reality (AR) and affective computing (AC). Furthermore, 3 different scenarios on tele-collaboration by virtual co-location are proposed for experiments with AC-AR studies. A system architecture is presented as support for the development of AC-AR tele-collaboration applications enabling joint work of local workers and remote experts.

Keywords

Affective Computing, Tele-collaboration, Augmented Reality.

Classification ACM

K.4.3 Organizational Impacts: Computer-supported collaborative work; H.5.1 Multimedia Information Systems: Artificial, augmented, and virtual realities.

INTRODUCTION

In modern work environments, even more complex technologies, protocols and scenarios are required not only to increase the productivity of workers, but also to fulfill basic work tasks. Often, typical work scenarios demand for mixed teams of workers that have different levels and types of expertise [56]. Either we talk about real work scenarios or just training/simulation sessions, inconsistency attributed to human factor or equipment typically rises serious problems further leading to the temporary inability to perform optimally the assigned tasks. Such problems may refer to situations when:

- The documentation is not sufficient/complete,
- The expertise is not ready on time/on the spot,
- The complexity of the problem/solution restricts the transfer of knowledge between the local worker and a potential remote expert using standard means of communications (e.g. audio channels by mobile phones),
- The activities are conducted under impeding affective conditions associated to stress, tiredness, anger, un-vigilance, etc.

The negative impact dimension related to the aforementioned situations increases exponentially for critical operations executed in specific work domains for which failure means the loss of equipment, property and even life. This applies especially to space missions, medical environments and to security-oriented operations (military, police, fire brigade, etc.). In this context, it becomes more important the urge to engineer systems that:

- Enable seamless collaboration among team workers,
- Automatically sense and adapt to the workers' state.

Current technology already permits access to partly or even completely understanding of behavior, intent and environment of a person, by automatic computer (software – SW) systems. As clearly indicated by Kanade and Hebert [23], the first-person vision (FPV) represents the most optimal way to sense the environment and the subject's activities from a wearable sensor. As compared to the more traditional approach of placing stationary sensing devices typically used in surveillance domain, better results concerning the automatic assessment can be obtained with data taken from the subject's environment, as sensed from his view point. From this perspective, the use of light weight and more powerful Head Mounted Device (HMD) having attached video cameras as well as other sensors, requires more attention, this having the potential to become a standard equipment in many work environments.

Augmented Reality (AR) technology already proved to have a significant impact in various domains [5]. Due to the capability to enhance reality, to assist collaboration, to support spatial cues and to allow interaction between the virtual and augmented worlds, AR promises to successfully enhance novel types of interfaces for face-to-face and remote collaboration [3].

Tightly related to AR, the concept of virtual co-location [7] implies the creation of spaces in which people and objects are either virtually or physically present: it allows people to engage in spatial remote collaboration. Virtual co-location entails that people are virtually present at any place of the world and interact with others that are physically present in another location to solve complex problems as if being there in person. State of the art research approaches on collaboration by virtual co-location in AR have showed considerable success for learning, simulations and professional practice, in mining industry [1], (serious) gaming [7][8][15][40] architecture [19][26], crime-scene investigation [11][31], space [6], crisis management [28], design [39][19], engineering [22], maintenance [17], weather [37], coordination [34].

The paper presents a AC closed-loop adaptive tele-collaboration AR system that is based on the DistributEd CoLlaborative AR Environment (DECLARE), a platform that supports virtual co-location of the local worker and the remote expert, and the transfer of knowledge by marker-less of expert using a state-of-the-art robust SLAM technique [36].

In order for a SW system to be adaptive with regard to the workers, first this must be able to sense the workers' state. An essential characteristic of the workers' state genuinely refers to emotions. Affective Computing (AC) is a technology that "relates to, arises from, or deliberately influences emotions" [32].

Even though considerable research approaches have been proposed separately on AR and AC, up to date there is no solid base for a systematic integration of the AC into the AR-based tele-collaboration process.

This paper describes an adaptive AC tele-collaboration AR system built on top of DECLARE platform and that follows the closed-loop AC system proposed by Wu et al. [41]. The system consists of three components namely the affect recognition component, the affect-modeling component and the affect control component. The affect recognition component performs the assessment of the user's affect by analyzing different body signals such as facial expressions, emotion in speech, psychophysiological signals, etc. The second component, the affect modeling creates a relationship between the user's affect and the features of the user's environment. The third component, the affect control provides the means for adapting the environment in such a way to get the user to the target affect state.

The automatic sensing, modeling and control components that regulate the system adaptation to the workers' state consider the subject's performance as the main criterion for the AR tele-collaboration experience.

According to the Yerkes-Dodson Law, the subjects' performance in mental tasks is dependent on the arousal in the form of non-monotonic function (Figure 1). Performance increases with arousal, given arousal is at low levels, reaches the peak at a given arousal level and decreases after that optimal level.

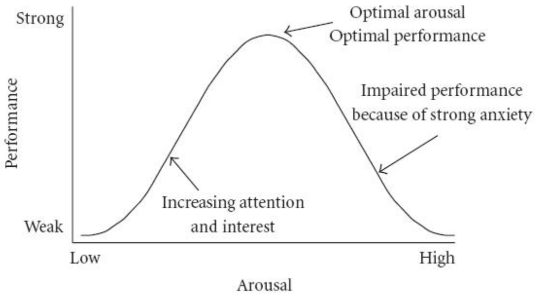


Figure 1. The Hebbian version of the Yerkes Dodson Law (Wikipedia)

This paper addresses the scenario analysis and the requirements of both theoretical and practical aspects that are necessary to design an AC-enabled tele-collaboration AR system for training and real work scenarios.

The rest of the paper is organized as follows: next section describes the tele-collaboration AR scenario and the design of the three components of the closed-loop AC system. The next section presents the cases for testing the AC tele-collaboration AR system. A further section details the design of DECLARE, a SW framework that facilitates easy development of AC-enabled tele-collaborative AR applications. Then, methodology for evaluation is discussed. Finally, conclusions and future work are presented.

TELE-COLLABORATION IN AR SCENARIO

The basic tele-collaboration AR scenario implies one local user and one remote expert collaborate by using an AR support system. Depending on the complexity of the collaborative work, variations to the basic scenario may include several local users or several remote experts.

The local user wears a HMD for AR which is connected to a portable computer. Additional (AC) body sensors are used to automatically assess the state of the local user.

The process of tele-collaboration in AR assumes both the local user and the remote expert are virtually co-located, both accessing the shared visual representation of the physical environment sensed through the video camera of the local worker's HMD. In the AR system, the virtual co-location is realized by transferring the view of the local user to the laptop computer of the remote user and by augmenting the shared view with inputs from the remote. In this way, the visualization on both sides is centered on the physical place of the local user.

The user interfaces for both local and remote users are designed following the usability principles of Nielsen and Molich [27].

Given the shared live video stream, the remote provides just-in-time and just-in-place assistance, based on his direct observation of the local user's activity.

For the local worker, visual displays of the sensed physical world together with additional information are furnished as augmented visual content through the HMD view.

Besides using the audio communication, the remote user can interact with the AR system running on his laptop computer to produce AR content which is shared with the local user. This is done in two ways, by adding augmented content and by altering representations of the sensed physical objects.

In addition to visualizing the video stream from the first person vision on local user's view, the remote user gets the 3D map of the local's physical environment. The insertion of AR content implies an authoring process to add virtual tags, attention-directing symbols including pointers, pop-up balloon like on-screen text messages, image representations and video content, static and animated models of 2D and 3D objects. Altering the visual representation of existing physical parts has the role to increase the level of attention towards the physical elements which are relevant for the assigned task.

Two scenarios are considered for AC-AR based tele-collaboration, given the type of system interface for the remote user.

Scenario 1

The AR system for the remote expert runs on a laptop or desktop computer. The input is based on standard mouse and keyboard devices, and the output is projected directly on the computer display. The tele-collaboration in AR scenario is illustrated in Figure 2. A special screen-based user interface enables the remote user to handle content in the AR overlay.

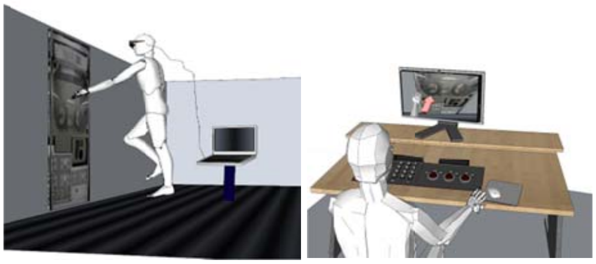


Figure 2. Tele-collaboration scenario with a local worker (left) and a remote expert using the desktop computer (right)

Scenario 2

A second scenario implies the remote expert can visualize the shared augmented world through the VR Oculus Rift, instead of on the screen of his laptop or desktop computer (Figure 3). Oculus Rift [49] (Figure 4; d) is an HMD for VR that has a resolution of 640×800 and the field of view of 110deg.

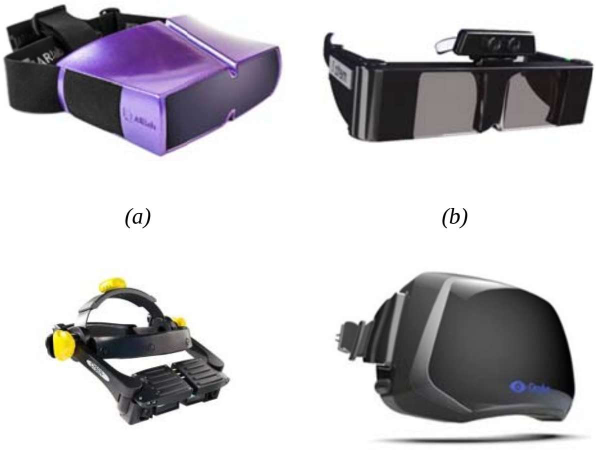
As in scenario 1, the remote user is basically coupled to the view of the local user. There are two possibilities for the remote to access the virtual environment, namely to stay coupled to the local's view or to freely navigate in the environment. The free navigation in the VR environment may be beneficial, the decoupling from the local's view and synchronizing the system visual output with the remote's own gestures providing the correct feedback for the user.

The interaction for the remote is by hand gesture and can be supported through computer-vision methods [9] or by using more accurate hardware devices such as Myo [51], Leap Motion [50] or SoftKinetic [52] sensors. Myo is a gesture control armband, Leap Motion and SoftKinetic are 3D depth sensors.

Figure 3. Tele-collaboration scenario with a local worker (left) and a remote expert wearing VR HMD Oculus Rift (right)

AR equipment

The system provides support for one type of video see-through HMD (MARTY - Figure 4a) and two types of optical see-through HMDs (Figure 4b,c). The open source HMD Marty [46] consists of a SONY HMZ-T1 optics based VR HMD modified in order to support two Logitech C905 HD webcams. A special 3D printed plastic case replaces the original SONY case of the headset. The first optical see-though HMD supported is META SpaceGlasses ver.1 [48] (960 x 540 transparent TFT LCD displays, a 720p RGB camera, infra-red depth camera and a 9 degree sensor). The second optical see-through HMD is Cyber-I SXGA Bi-ocular HMD [47] (SXGA, FOV 50deg).



(a) *(b)*

(c) *(d)*

Figure 4. AR HMDs: MARTY (a), META SpaceGlasses ver.1 (b), Cyber-I SXGA Bi-occular HMD (c) and Virtual Reality Oculus Rift (d)

The video streams acquired from HMD's stereo camera, are used in the process of preparing the AR view and as input for the automatic sensing modules (i.e. for the computer-based hand detection module). In addition to video data, some HMDs provide extra information. The META SpaceGlasses ver.1 HMD has attached an infrared depth sensor (used for sensing the physical environment and user's hands). Cyber-I HMD can have attached two eye-trackers, one for each eye.

Affective Computing system

The AC systems proposed so far mostly approach the recognition of the affect alone. A few studies have been conducted on AC closed-loop architectures that use the affect information to prepare the system feedback and to control the user experience. Omata et al. [29] describe a system that can influence a viewer's arousal by adapting visual parameters such as color and shape. Yannakakis [38] presents a real-time adaptation mechanism that controls game parameters in order to optimize player satisfaction in AR games. Toups et al. [35] finds that tracking attention, effort and stress by analyzing electrodermal and electromyographic activity may lead to new game strategies, greater level of attention and new forms of focused involvement in team game playing. Parsons and Reinebold [30] describe a VR framework that use neurocognitive and psychophysiological estimations to dynamically adapt the difficulty level in serious games, aiming to increase the readiness towards return-to-duty state for soldier patients. Wu et al.[41] make use of affective and cognitive assessments as part of the design for adaptive environments, with the goal of getting the user to the optimal performance state.

In the context of tele-collaboration in AR, an AC closed-loop design is adopted to automatically assess, model and control the workers' affect and performance during the execution of specific team-oriented tasks, where local workers are remotely supported by remote experts. The AC system has 3 components, namely for the recognition of the affect state, the modeling and the control. The AC components are further discussed.

Recognition of worker's state

A fundamental characteristic of the user's state is the affect. A basic model in referencing the affective space (circumplex model of emotion of Russel [33]) relates to two dimensions, the valence and the arousal of the emotion. The valence represents the hedonic value: positive/negative and the arousal represents the intensity of the emotion.

To characterize the worker's state, an estimation of his/her affect is realized by considering the two dimensional model of emotions as the encoding scheme. The affect can then be estimated by processing body signals and extracting specific measures such as psychophysiological indicators, facial expressions, emotion in speech (prosody, semantics), body gestures, etc.

Psychophysiological computing relies on direct measurements of the activity of the brain [21][41] and indirect measurements of the cardiorespiratory activity [29][30][41], the electro-dermal activity [29][30][35][41], respiration [29][30][41], pupillometry [30][41] or electromyography [35]. The cardiorespiratory activity may consider the heart rate – HR or the heart rate variability – HRV. The electro-dermal activity may consider the skin conductance or galvanic skin response – GSR.

Figure 5. Physiological body sensors [45] using ARDUINO [44]

In this research, the analysis of physiological body signals is run using e-Health Sensor (Figure 5) [45], a HW equipment which is supported by ARDUINO platform [44], and Cortrium sensors [55]. The following sensors are considered:

- Position Sensor (Accelerometer)
- Body Temperature Sensor
- Blood Pressure Sensor (Sphygmomanometer)
- Pulse and Oxygen in Blood Sensor (SPO2)
- Airflow Sensor (Breathing)
- Galvanic Skin Response Sensor (GSR - Sweating)
- Electrocardiogram Sensor (ECG)
- Electromyography Sensor (EMG)

In addition, the estimation on the worker's state makes use of automatic models for facial expression recognition [12] (Figure 6), non-contact heart rate detection [10] and point-of-interest analysis, by using eye trackers.

Figure 6. Automatic Multimodal Emotion Recognition prototype by facial expressions and emotion clues in speech [12]

Figure 7 illustrates a snapshot of a SW system prototype that runs non-contact automatic heart rate detection, based on the analysis of the face skin color variability. More accurate estimations of the heart rate are further obtained by fusing this result with the indications by other physiological sensors.

Figure 7. Non-contact automatic heart rate detection system prototype by face analysis [10]

Figure 8 illustrates a snapshot of a system prototype for affect recognition which relies on data fusion models that analyze body gestures and physiological signals. In the figure, the local worker is a police officer who wears an optical-though HMD (META) and an AR mobile system assembled in a backpack.

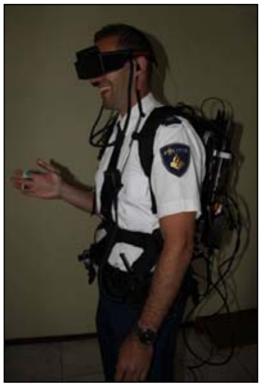


Figure 8. Police officer equipped with the experimental AC-AR kit (META HMD [48] and e-Health Sensor [45])

Modeling

The affect modeling component holds a relationship between the characteristics of the tele-collaboration AR environment and the dynamics of worker's affective state. Relevant data is collected during an experiment session in which the workers perform domain specific tasks collaboratively, by using AR technology.

The model is generated during the training phase of a supervised learning algorithm which runs either off-line - after the experiment session, or on-line. Once the learning phase is completed, the model takes the form of a knowledge dataset which encodes the mapping from the affect states to the control vectors that represent variations in the parameter space that define the tele-collaboration AR environment.

Control

The control component outputs a control signal which alters the parameters of the tele-collaboration environment and so stimulates the worker's affective state to change into a desired direction. The change is made so that the performance of the worker using the AR system improves.

CASES OF TELE-COLLABORATION IN AR

Several scenarios have been designed for conducting research on the AC technology and the tele-collaboration by AR. The AC technology has the role of sensing the workers' affective state and to adjust the system based and tele-collaboration based parameters accordingly, so as that the team performance is optimal.

The remote experts can actively augment the shared view by adding relevant visual content in form of text, photos, videos, static and animated objects pointing to physical locations being currently surveyed and being closely related to the current task to fulfill.

Ground training for In-Flight Maintenance onboard ISS Columbus

The International Space Station (ISS) [42] is a modular structure that provides a platform to accommodate several scientists from different fields conducting scientific research in space (Figure 8). In-Flight Maintenance (IFM) operations on-board of ISS Columbus imply astronauts fulfill specific quantifiable tasks regarding setting up, checking, repairing the equipment, and conducting mission specific procedures such as those involving the completion of scientific experiments.

Figure 8. Researcher wearing an AR HMD onboard ISS Columbus mock-up

The intrinsic nature of the equipment and procedures makes all manned spaceflight operations exhibit a high degree of complexity. The crew fundamentally lacks in-depth expertise on all on-board systems, critical situations typically being handled with support from local sources of information and from remote specialists located at the ground base.

The main local source of information is a laptop computer, the Station Support Computer (SSC) which stores complete information about all standardized operations including IFM. To clarify a task, an astronaut has to discuss with on-board members, with members located at the ground base and to access SSC, every time temporarily stopping his current activity. Shifting the attention to other task possibly related to a slightly different physical location may distract and increase the workload to the astronaut. Furthermore, the unavailability of external immediate support in case of emergency situations requires the missions to be carried on with extra caution and under the lowest possible error rate of human operation.

The objective of the research [6] is to focus on novel ways to use AR technologies in the human space flight domain, with emphasis on space collaborative scenarios as well as on remote support in space operations. AR in conjunction with the use of HMDs and additional collaboration technologies, are intended to provide the base for improving the collaboration and the remote support on hand-free operations, implementing a just-in-time and just-in-place assistance approach for training and ground simulations (Figure 9). Even more, the technologies are envisioned to narrow the gap between training and real operations as the very same technology is expected to be used in both contexts without modification.

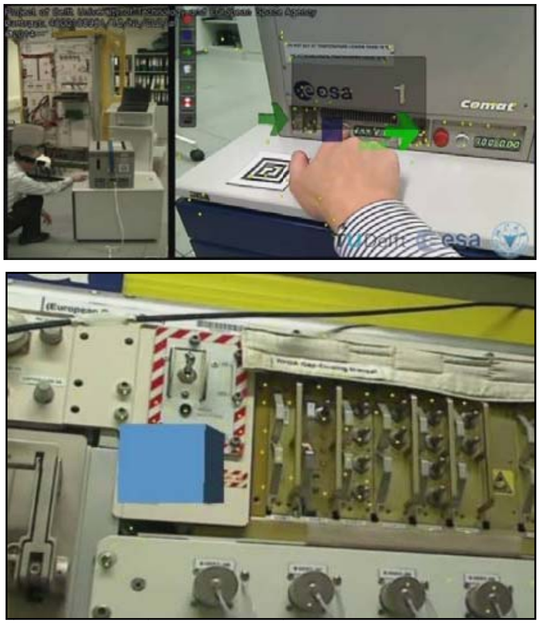


Figure 9. Examples of Augmented Views for the astronaut/trainee onboard ISS Columbus mockup

Security

For operational units in the security domain that work together in teams, it is important to quickly and adequately exchange context-related information. Wrong decisions or choices may affect the continuation of an operation, the security of the operational units as well as the possibly affected civilians. AR systems can help operational security teams by providing local support for detection, recognition and reconstruction in a physical environment (Figure 10), and by facilitating the collaboration with remote experts.

Figure 10. Automatic object detection – gun (top left), marker-based AR bloodstain pattern reconstruction (top right) and automatic face recognition (down)

A previous project with the Netherlands Forensic Institute proposed a novel mediated reality system for collaborative spatial analysis on location [11][31][43].

At the moment, a joint project with security institutions represented by the Dutch Police and the Netherlands Forensic Institute (NFI) investigates the potential of AR techniques to facilitate information exchange and situational awareness of teams in the security domain.

Three different scenarios from the security domain have been elicited using an end-user oriented design approach. The 3 identified scenarios are:

- **Reconnaissance teams**: A policeman, equipped with a head mounted device (HMD) investigates a safe house in which a witness needs to be safely accommodated. This policeman shares the local view as recorded from the HMD camera with a remote colleague. While the local policeman investigates the safe house, the remote agent has the task to highlight suspect objects in the house and point out possible emergency exits.
- **Forensic investigation**: A forensic investigator arrives at a severe crime scene. Wearing an HMD, the investigator shares the local view with a remote colleague. The remote colleague has the task to point the local colleague to possible evidence, take pictures of evidence and support the preparation of 3D laser scans.
- **Domestic violence**: A team of 2 policemen arrives at a scene of domestic violence. One of the policemen wears an HMD and shares the local view with a remote colleague. The remote colleague can provide instructions and information on the case, take pictures and highlight possible evidence. The local policeman wearing the HMD orally shares received information with the second local colleague.

The user interface, created using Unity 3D game engine [54], was customized according to the requirements of each scenario and was adjusted to the role of each player. The user interface design follows the usability principles of Nielsen and Molich [27].

The view of the local person is adapted to the used HMD (optical see-through META HMD, Figure 4b), and shows only virtual objects currently visible in the grey area illustrated in Figure 11 (right side). For the remote user, the interface displays the live video captured from the local's HMD camera. Additionally, it shows an authoring menu with buttons for handling virtual 2D and 3D objects.

Figure 11. Setup for conducting experiments on AC-based tele-collaboration in AR for safety domain

For instance, in the crime investigation scenario (Figure 11), the remote person is able to place 3D objects (spheres, cubes, arrows), to write 3D text messages, to place laser stickers to mark physical areas to be scanned by the local investigator, and to take/load photos from the local scene. The transparent rectangular region in the middle of the screen (Figure 11, right displays) represents the view being shared by both remote and local investigators.

Based on the scenarios, an experiment took place at a training location of SWAT teams, in which 11 policemen and inspectors from 4 operational units of 3 national Dutch security institutions participated. An usability study was conducted based on the experiment.

AR Tower Game

A collaborative game has been designed to approximate collaborative solving of shared complex problems and to explore the different perception of presence in AR scenarios [7][8]. The game requires players to share their expertise to collaboratively build a tower with differently colored physical blocks or virtual blocks in AR. In the game, the workers are replaced with team players collaboratively building a tower in AR. The game can be adjusted dynamically by tracking the affective state of the players using AC technology.

The goal of the AR game is to jointly build a tower by using the colored blocks available on the game board. The game represents an approximation of a complex shared task. The rules of the game are made in such a way to support collaboration among participants. Partial information available at any time allows for nonlinear game dynamics, process of discovery and gradual development toward fulfilling the shared game task, the building of the tower.

Individual expertise is represented by the ability of each player to manipulate blocks of specific color only. Shared expertise is represented by the ability of all players to move blocks of the same color. The order of the blocks in the tower to be built is not to be randomly set. Instead, the color order of the blocks has to contain the individual color pattern of each user. This color pattern represents the individual task of the players. The individual expertise and the individual task reflect the knowledge of the players and are defined at the beginning of the game session. The shared goal of building the tower is then achieved through a sequential process in which the players have to communicate and to agree upon the action strategy involving the next block to be moved. None of the players can build the minimal tower containing the individual color pattern without help from at least one of the other players.

The game session implies three players, two of them being collocated at the ‘construction site’ and a third one being located in a separate room (Figure 12). Each physically collocated player sits at the table and wears a stereo AR HMD, with direct view of a single AR pattern marker placed on the table.

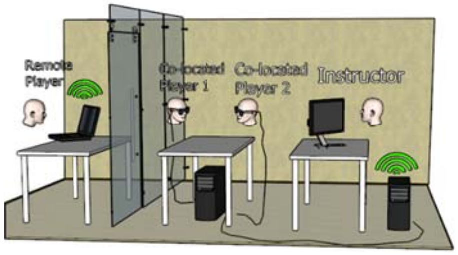


Figure 12. Experimental setup for the AR game

The AR system of physically collocated players projects the augmented game elements as an additional layer on top of the stereo camera video stream, into the AR stereo view of the users. The block construction site is centered at the position of the AR pattern (Figure 13).

Figure 13. Testing scenario for the tele-collaborative AR game

The collocated players access the AR game system by free hand interaction (Figure 14). The system supports multiple inputs from mouse devices, AR pattern in the form of ‘magic wand’, color pattern for the tip of the finger and palm posture detection [9]. An augmented cursor (blue circle) is shown at the location where the tip of the finger is detected. Compared to the co-located users, the remote user interacts by using a regular mouse device.

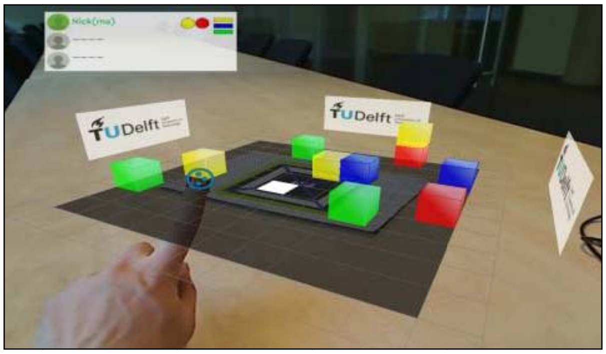


Figure 14. Interaction by free-hands while playing AR game

During the game, the system tracks the actions of players against consistency of rules with individual goal assignments. An on-screen message is then displayed every time a player attempts to select a block of a color other than the colors defining his/her abilities.

DISTRIBUTED COLLABORATIVE AUGMENTED REALITY ENVIRONMENT (DECLARE)

The above-described scenarios have been supported by using a customized framework – DECLARE, an extension of a previous architecture for automatic surveillance applications [14]. This is based on the shared memory space which decouples data communication in both time and space.

The framework integrates a set of functional modules which provide support for two different categories of systems, for the local and for the remote users. Each category of software system runs a Unity-based [54] customized user interface which is adapted to the interaction needs of the user.

The framework supports virtual co-location of multiple users simultaneously, through specialized applications serving both local and remote roles. The communication is enabled through a shared memory mechanism which decouples data transfer in time and space. The architecture is scalable, distributed and modular, with a range of functional modules e.g. hand tracking and gesture recognition [9].

An essential part of DECLARE is represented by the RDSLAM module which is based on the state-of-the-art marker-less algorithm for mapping the physical environment and for localizing the user in the environment [36]. The sparse cloud of 3D points generated by RDSLAM further provides virtual anchors between the AR and the physical environments that enable AR annotations by the remote user. DECLARE integrates functionality for affective computing (“AC +Sensors” box) for both remote’s and local user’s systems.

The diagram in Figure 15 suggests a functional integration of modules ranging from the sensory low level to the high semantic level, in the same time by considering static task information as well as dynamic aspects related to the interaction between the local and the remote user. The modules run specific tasks related to network communication, data logging services, interaction management and the implementation of complex video processing algorithms.

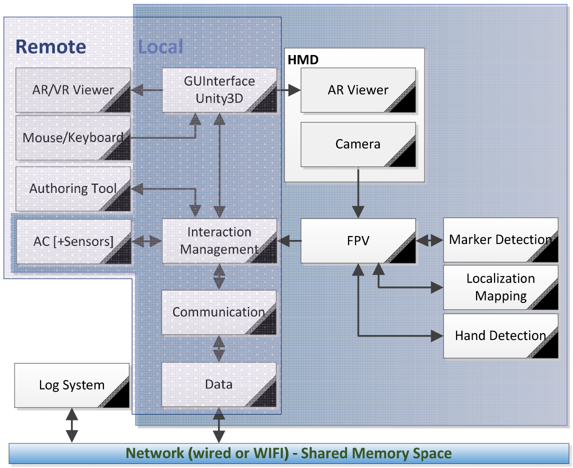


Figure 15. System diagram of DECLARE's functional modules

Local User

For the local user, at the sensory level the input consists of the video stream collected from the video camera attached to the HMD of the user ("Camera" box). The raw data can be seen as the input to the First Person Vision module ("FPV" box). According to the work context, the FPV module proceeds with the automatic processing of the input video data using one of the available modules for detecting the physical markers for AR ("Marker Detection" box), for localizing and mapping the physical environment ("Localizing Mapping" box) or for detecting the user's hands ("Hand Detection" box). Subsequent to the automatic processing of the input video data, the "FPV" module passes the results related to the 3D position and orientation of the AR marker, the 3D location and orientation of the trainee (by approximation from the position of her HMD) and the location and possibly appearance of the local users' hands in the image. All the information is passed to the interaction management module ("Interaction Management" box).

The robustness at the data communication level assumes that the system running on the local user's side is still able to function even though the network connection with the remote system is interrupted. In such a situation, the local user can still visualize the video content of the stereo camera attached to his HMD and benefits from the information extracted automatically by the local data processing modules.

The network communication failures denote a critical issue, especially for the case of real work contexts making use of wireless data networks.

Remote user

On the remote's side, the input consists of the raw video stream provided by the local user's system via the shared memory space, together with the high-level, semantic information related to the position of AR markers, user's hands and user's location and orientation in the physical world.

The transfer of the combined low-level video data (from the HMD camera of the local user) together with high level semantic information of the local user's interaction represents the synchronization mechanism which enables the virtual co-location process. The system for the remote makes use of a user interface which is adapted for the regular computer screen as output device and the mouse and keyboard as input devices (Figure 2). The second possibility for the remote user is to use an HMD for VR (Figure 3).

The output for the user interface on the system for the remote user is handled through the GUI interface module ("GUInterface Unity3D" module) and an AR/VR module ("AR/VR Viewer" box). The input is handled by a mouse/keyboard module ("Mouse/Keyboard" box). The GUI for the remote user supports an authoring tool ("Authoring Tool" box) that has functionality for adding and removing augmented content on the AR view which is shared between local and remote users. Similar to the case of the AR system for the local user, the AR system for the remote user integrates an interaction management module ("Interaction Management" box). This module receives as input both low-level video data and high-level semantic data from the AR system of the local (via the shared memory space). Additional inputs to the module are the remote's interaction events received from the authoring tool and static task-oriented descriptions. Based on these inputs, the interaction management module generates the output to the AR/VR module which, in turn formats the output according to the hardware devices used for visualization (standard computer screen or VR HMD).

EVALUATION METHODOLOGY

A series of experiments were already conducted on studying the usability of the tele-collaboration AR systems specially designed to support remote and local workers in the above-described scenarios. These systems were developed using DECLARE framework.

A new goal is to study the efficiency of the tele-collaboration systems which integrate affective computing (AC) technology. The study may explore the usability of the tele-collaboration systems in two conditions, with and without AC support.

The evaluation considers both subjective and objective measurements from the data collected during specific experiments with regard to the system usability and performance of the tasks to be accomplished by the virtually-collocated team.

Questionnaires, semi-structured interview and the expert analysis on the video recordings of the participants during experiments, are appropriate instruments providing valuable subjective indicators. As objective measurements, the analysis may take into account system logs from the participant's interaction with the user interface, and results from automatic assessment of body behavior (face, hand and body gestures; AC measurements).

The questionnaires may consider the measurement of AC-AR system usability by using the Post-Study System Usability Questionnaire (PSSUQ) [25] and the System Usability Scale (SUS) [4]. In addition, the questionnaires may include the NASA Task Load Index (TLX) [20], the AR presence questionnaire of Gandy et al. [18] and the situational awareness questionnaire of Endsley [16]. The questions relate to the assessment of the usability, workload, interaction, interface, tactile experience, moving in the environment and the measurement of situational awareness.

CONCLUSIONS AND FUTURE WORK

The paper presented an approach to integrate two technologies, namely affective computing and tele-collaboration in augmented reality. A system architecture is proposed as support for the development of AC-AR tele-collaboration applications enabling joint work of local workers and remote experts. Furthermore, 3 different scenarios were described for which several studies have been already conducted on tele-collaboration by virtual co-location. As future work, the AC-AR system prototypes will be evaluated based on the scenarios described.

ACKNOWLEDGEMENTS

The work for the current paper has been done as part of a joint project with The European Space Agency, the ESA Contract number 4000108961/13/NL/GLC/al, and of a joint project with Dutch National Safety Institutions.

REFERENCES

1. Alem, Leila; Tecchia, Franco, Huang, Weidong. Remote Tele-assistance System for Maintenance Operators in Mines, *11th Underground Coal Operators' Conference*, 171-177 2011.
2. Billinghurst, Mark ; Weghorst, Suzanne and Furness, Tom, III. “Wearable computers for three dimensional CSCW”, *First International Symposium on Wearable Computers, Digest of Papers*, 39 – 46, 1997.
3. Billinghurst, Mark and Kato, Hirokazu. “Collaborative augmented reality”, *Communications of the ACM - How the virtual inspires the real CACM, 45(7)*, 64-70, 2002.
4. J. Brooke, J., “SUS: a "quick and dirty" usability scale”. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland. *Usability Evaluation in Industry*. London: Taylor and Francis, 1996.
5. Carmigniani, Julie; Furht, Borko; Anisetti, Marco; Ceravolo, Paolo; Damiani, Ernesto and Ivkovic, Misa. Augmented reality technologies, systems and applications, *Multimedia Tools and Applications, 51:1*, 341-377, 2011.
6. Dragos Datcu, Marina Cidota, Stephan Lukosch, David Martinez, Mikael Wolf, Virtual Co-location to Support Remote Assistance for Inflight Maintenance in Ground Training for Space Missions, *ACM CompSysTech'14*, Ruse, Bulgaria, 27 June 2014.
7. Dragos Datcu, Stephan Lukosch, Heide Lukosch, A Software System for Studying the Perception of Presence in a Collaborative Real World and Augmented Reality Scenario, *Proc. CSCW 2014*, Baltimore, 2014.
8. Dragos Datcu, Stephan Lukosch, Heide Lukosch. Comparing Presence, Workload and Situational Awareness in a Collaborative Real World and Augmented Reality Scenario, *IEEE ISMAR 2013* Workshop on Collaboration in Merging Realities (CiMeR), Adelaide, SA, Australia, October 1, 2013.
9. D. Datcu, and S. Lukosch, “Free-Hands Interaction In Augmented Reality”, *Proceedings of the ACM Symposium on Spatial User Interaction (SUI)*, July 20-21, Los Angeles, CA, 2013.
10. Dragos Datcu, Marina Cidota, Stephan Lukosch, Leon Rothkrantz, Noncontact Automatic Heart Rate Analysis in Visible Spectrum by Specific Face Regions, *International Conference on Computer Systems and Technologies – ACM CompSysTech'13*, Ruse, Bulgaria, 28-29 June 2013.
11. D. Datcu, T. Swart, S. Lukosch, and Z. Rusak, Multimodal Collaboration for Crime Scene Investigation in Mediated Reality. *Proceedings of the ACM International Conference on Multimodal Interaction (ICMI)*, Santa Monica, California, USA, October 22-26, 2012.
12. D.Datcu, “Multimodal Recognition of Emotions”, PhD. Thesis, Delft University of Technology, 2009.
13. Dragos Datcu, Leon J.M.Rothkrantz, "Automatic bi-modal emotion recognition system based on fusion of facial expressions and emotion extraction from speech ", *IEEE Face and Gesture Conference FG2008*, Amsterdam, The Netherlands, 17-19 September 2008.
14. Datcu, D., Yang, Z., Rothkrantz, L. Multimodal workbench for automatic surveillance applications, *IEEE Workshop CVPR 2007*, Minneapolis, 2007.
15. Dong, Suyang; Behzadan, Amir H.; Chena, Feng and Kamat, Vineet R. Collaborative visualization of engineering processes using tabletop augmented reality, *Advances in Engineering Software 55*, 45-55, 2013.
16. Endsley, M. R. Toward a Theory of Situation Awareness in Dynamic Systems, *Human Factors and Ergonomics Society, 37*: 32-64, 1995.
17. Ferrise, Francesco; Caruso, Giandomenico and Bordegoni, Monica. Multimodal training and tele-assistance systems for the maintenance of industrial products, *Virtual and Physical Prototyping, 8:2*, 113-126, 2013.
18. Gandy, M.; Catrambone, R., MacIntyre, B., Alvarez, C., Eiriksdottir, E., Hilimire, M., Davidson, B., and McLaughlin, A.C. “Experiences with an AR evaluation test bed: Presence, performance, and physiological measurement”, *Proceedings of ISMAR*, 127-136, 2010.
19. Gu, Ning; Kim, Mi Jeong and Maher, Mary Lou. “Technological advancements in synchronous collaboration: The effect of 3D virtual worlds and tangible user interfaces on architectural design”, *Automation in Construction, 20*: 3, 270-278, 2011.
20. Hart, S. “Nasa-Task Load Index (Nasa-TLX); 20 Years Later”, P*roceedings of HFES 50th Annual Meeting*, 904-908, 2006.

21. Robert Horlings, Dragos Datcu, Leon J.M. Rothkrantz, "Emotion Recognition using Brain Activity", CompSysTech'08, ISBN 978-954-9641-52-3, *ACM International Conference Proceedings Series, no. 374*, July 2008.
22. Jailly, Benjamin; Gravier, Christophe; Preda, Marius and Fayolle, Jacques. “Interactive mixed reality for collaborative remote laboratories”, *Proceedings of the third international ACM workshop on Multimedia technologies for distance learning - MTDL '11*, ACM New York, NY, USA, 1-6, 2011.
23. Kanade, T. and Hebert, M. “First-Person Vision”, *Proceedings of the IEEE, Volume 100*, Issue 8, 2442-2453, 2012.
24. G. Klein, D. Murray, Parallel Tracking and Mapping for Small AR Workspaces, *Proc. International Symposium on Mixed and Augmented Reality*, 225-234, 2007
25. J.R. Lewis, “Psychometric evaluation of the post-study system usability questionnaire: the PSSUQ", *Human Factors Society*, 1259-1263, 1992.
26. Moeslund, T. B.; Störring, M.; Broll, W.; Aish, F.; Liu, Y. and Granum, E. “The ARTHUR System: An Augmented Round Table”, *Journal of Virtual Reality and Broadcasting, 34*, 2003.
27. Nielsen, J., and Molich, R, “Heuristic evaluation of user interfaces”, *Proc. ACM CHI'90* (Seattle, WA, 1-5 April), 249-256, 1990.
28. Nilsson, Susanna; Johansson, Bjorn and Jonsson, Arne. “Using AR to support cross-organisational collaboration in dynamic tasks”, *Proc. ISMAR 2009*, 3-12, 2009.
29. Omata, M.; Moriwaki, K.; Xiaoyang Mao; Kanuka, D.; Imamiya, A., "Affective rendering: Visual effect animations for affecting user arousal," *Multimedia Computing and Systems (ICMCS), 2012 International Conference on* , vol., no., pp.737-742, 10-12 May 2012.
30. Parsons, Thomas; Reinebold, James. ““Neuroscience and Simulation Interface for Adaptive Assessment in Serious Games” *IEEE International Games Innovation Conference - IGIC*, 93-96, 2011.
31. Poelman, R., Akman, O., Lukosch, S., & Jonker, P. As if Being There: Mediated Reality for Crime Scene Investigation. *Proc. CSCW’12*, 1267–1276, 2012.
32. Rosalind W. Picard. *Affective Computing*, MIT Press, 1997.
33. Russell, James, "A circumplex model of affect". Journal of Personality and Social Psychology 39: 1161–1178, 1980.
34. Schnier, Christian; Pitsch, Karola; Dierker, Angelika and Hermann, Thomas. “Collaboration in Augmented Reality: How to establish coordination and joint attention?”, *Proc. CSCW 2011*, Aarhus Denmark, 405-416, 2011.
35. Zachary O. Toups, Ross Graeber, Andruid Kerne, Louis Tassinary, Sarah Berry, Kyle Overby, and Maritza Johnson. A Design for Using Physiological Signals to Affect Team Game Play. In *ACI '06: Proceedings of Augmented Cognition International*, Oct 2006.
36. Tan, Wei; Liu, Haomin; Dong, Zilong; Zhang, Guofeng and Bao, Hujun. “Robust Monocular SLAM in Dynamic Environments”, *International Symposium on Mixed and Augmented Reality ISMAR 2013*, Adelaide, Australia, 2013.
37. Yabuki, Nobuyoshi; Furubayashi, Shuhei; Hamada, Yuuki; Fukuda, Tomohiro. Collaborative Visualization of Environmental Simulation Result and Sensing Data Using Augmented Reality, *Cooperative Design, Visualization, and Engineering, Lecture Notes in Computer Science, 7467*, 227-230, 2012.
38. Yannakakis, G.N., “Real-time adaptation of augmented-reality games for optimizing player satisfaction”, Computational Intelligence and Games, 2008. CIG '08. *IEEE Symposium*, 103 – 110, 2008.
39. Wang, Xiangyu and Dunston, Phillip S. Comparative Effectiveness of Mixed Reality-Based Virtual Environments in Collaborative Design, *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Volume 41, Issue 3*, 284-296, 2011.
40. Wichert, Reiner. “Collaborative Gaming in a Mobile Augmented Reality Environment”, *EUROGRAPHICS - Ibero-American Symposium in Computer Graphics - SIACG 2002*, 31-37, 2002.
41. Wu, D.D.; Courtney, C.G.; Lance, B.J.; Narayanan, S.S.; Dawson, M.E.; Oie, K.S.; Parsons, T.D., "Optimal Arousal Identification and Classification for Affective Computing Using Physiological Signals: Virtual Reality Stroop Task," *Affective Computing, IEEE Transactions on , vol.1, no.2*, 109,118, July-Dec. 2010.
42. http://www.esa.int/
43. CSI The Hague, the Netherlands Forensic Institute, http://www.csithehague.com/.
44. ARDUINO, http://www.arduino.cc/.
45. E-Health Sensors, http://www.cooking-hacks.com/.
46. Marty AR headset by Studio Raar, http://arlab.nl/media/open-source-headset-marty-now-downloadable, (last accessed June 2013).
47. Cybermind, http://www.cybermindnl.com/
48. META SpaceGlasses, https://www.spaceglasses.com/.
49. Oculus Rift, http://www.oculusvr.com/.
50. https://www.leapmotion.com/
51. https://www.thalmic.com/en/myo/
52. http://www.softkinetic.com/
53. OpenCV computer vision library, http://opencv.willowgarage.com/wiki/.
54. Unity3D game engine. https://unity3d.com/
55. http://cortrium.com/
56. van Knippenberg D, De Dreu CK, Homan AC., “Work group diversity and group performance: an integrative model and research agenda”, *Journal of Applied Psychology, 89(6)*, 1008-1022, 2004.